

**Investigating candidate genes identified by
genome-wide studies of granulomatous diseases in
susceptibility to tuberculosis: *ANXA11* and the
CADM family**

by
Muneeb Salie

*Thesis presented in partial fulfilment of the requirements for the degree
Master of Medical Science (Human Genetics) at the University of
Stellenbosch*



Supervisor: Prof. Eileen Garner Hoal
Co-supervisor: Dr. Marlo Möller
Faculty of Health Sciences
Department of Biomedical Sciences

December 2010

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the owner of the copyright thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Abstract

The infectious disease tuberculosis (TB) remains the leading cause of death worldwide by a single infectious agent, despite significant advances in biomedical sciences. The idea that host genetics plays a role in the development of disease was proposed by Haldane in 1949. The observation that only 10% of immunocompetent individuals develop disease while others are able to successfully contain it, further suggests that host genetics plays an important role. TB is thus a complex disease, with the causative bacterium, *Mycobacterium tuberculosis*, host genetic factors and environment all contributing to the development of disease. To date several genes have been implicated in TB susceptibility, albeit with small effect.

Genome-wide association studies (GWAS) offer the means to identify novel susceptibility variants and pathways through their ability to interrogate polymorphisms throughout the genome without being limited by our understanding of the immune processes involved in TB infection and disease progression. TB and sarcoidosis are both granulomatous diseases, and we therefore hypothesized that the genes and their associated variants identified in recent GWAS conducted in West Africa for TB, and Germany for sarcoidosis, could alter susceptibility to TB in the South African Coloured (SAC) population. In the sarcoidosis GWAS, ANXA11 was shown to alter susceptibility to sarcoidosis; whereas in the TB GWAS, CADM1 was found to alter susceptibility to TB.

This study tested the association with TB of 16 polymorphisms in 5 potential TB host susceptibility genes in the SAC population. A well designed case-control study was employed, using the TaqMan[®] genotyping system to type the various polymorphisms. Any polymorphism that was found to be significantly associated with susceptibility to TB was then subjected to further analysis to determine the functional effect of the polymorphism. Promoter methylation patterns were also investigated in ANXA11 as another mechanism to elucidate its role in TB susceptibility.

A 3' UTR ANXA11 polymorphism was found to be strongly associated with susceptibility to TB, including 3 haplotypes. The gene expression analysis identified differential transcriptional levels between individual with the different genotypes, with individuals homozygous for the A-allele exhibiting a 1.2-fold increase in gene expression relative to those homozygous for the G-allele. Methylation analysis however found no differences between cases and controls. In addition, 16 novel polymorphisms were also identified, 15 of which occurred in the 3'UTR of ANXA11. The mechanism of action of ANXA11 in TB susceptibility is hypothesised to be in the area of endocytosis, autophagy or apoptosis.

A weak association was noted with one of the 5' UTR polymorphisms of CADM3, which did not hold up to further analysis in the GWAS study, and no functional work was therefore done.

This work facilitates our understanding of the role of host genetics in susceptibility to TB and adds to the growing amount of information available. Proper understanding of the role that host genetics plays in TB susceptibility could result in better treatment regimens and prediction of individuals who are at a greater risk of developing TB, a disease that still kills millions of individuals annually.

Opsomming

Tuberkulose is verantwoordelik vir meer sterftes as enige ander aansteeklike siekte, ten spyte van die voortuitgang wat die Biomediese Wetenskappe tans beleef. In 1949 het Haldane voorgestel dat die genetiese samestelling van die gasheer 'n rol speel in vatbaarheid vir aansteeklike siektes. Vir tuberkulose word hierdie aanname gesteun deur die feit dat slegs 10% van individue wat geïnfekteer word aktiewe simptome ontwikkel, terwyl 90% die siekte suksesvol sal afweer. Tuberkulose is dus 'n komplekse siekte wat veroorsaak word deur *Mycobacterium tuberculosis*, maar wat beïnvloed word deur genetiese sowel as omgewingsfaktore. Verskeie gene is al geïdentifiseer wat 'n rol speel in vatbaarheid vir tuberkulose, tog is hul invloed betreklik klein.

Genoom-wye assosiasiestudies (GWAS) bied unieke geleenthede vir die identifisering van nuwe polimorfismes wat genetiese vatbaarheid kan beïnvloed. Hierdie tegniek kan die hele genoom fynkam, sonder dat enige vooropgestelde idees oor die immuunrespons teen tuberkulose 'n invloed sal hê. Tuberkulose en sarkoïdose is albei siektes wat die vorming van granulomas veroorsaak. Verskeie gene met hul geassosieerde variante is geïdentifiseer in 'n onlangse GWAS, wat gefokus het op populasies in Wes-Afrika en Duitsland. Ons hipotese was dat die polimorfismes wat in hierdie studie geïdentifiseer is, 'n invloed kan hê op genetiese vatbaarheid vir TB in die Suid-Afrikaanse Kleurlingbevolking (SAK). Die sarkoïdose GWAS het bevind dat *ANXA11* vatbaarheid vir die siekte beïnvloed, terwyl *CADM1* in die tuberkulose GWAS geïdentifiseer is.

Die studie het die assosiasie tussen 16 variante en tuberkulose vatbaarheid ondersoek in die SAK populasie. Die variante strek oor 5 potensiële tuberkulose vatbaarheidsgene. Goedbeplande pasiënt-kontrole assosiasiestudies is gedoen en die polimorfismes is gegenotipeer deur gebruik te maak van die TaqMan[®] genotiperingsstelsel. Enige polimorfisme wat beduidend met tuberkulose geassosieer was, is verder geanaliseer om die moontlike funksionele invloed daarvan te bepaal. Promotormetileringspatrone van *ANXA11* is ook geanaliseer, om 'n addisionele meganisme in tuberkulose vatbaarheid te ondersoek.

Na genotipering van die polimorfismes is 'n 3' UTR *ANXA11* variant geïdentifiseer wat beduidend met tuberkulose vatbaarheid geassosieer was. Drie haplotipes is ook geïdentifiseer. 'n Geenuitdrukkinganalise het aangedui dat verskille in transkripsie vlakke voorkom in individue met verskillende genotipes. Individue wat homosigoties was vir die A-alleel het 'n verhoging van 1.2 in geenuitdrukking gehad, relatief tot individue wat homosigoties was vir die G-alleel. Metileringsanalise het egter geen verskil aangedui tussen pasiënte en kontroles nie. Addisioneel, is 16 nuwe variante ontdek, waarvan 15 in die 3'UTR van *ANXA11* geleë was. Die meganisme waarmee *ANXA11* genetiese vatbaarheid vir tuberkulose beïnvloed, blyk in die area van endositose, apoptose of outofagie, te wees.

'n Swak assosiasie is gevind vir 'n 5' UTR variant van *CADM3* en is nie verder opgevolg in die GWAS nie. Gevolglik is geen funksionele studies op hierdie polimorfisme gedoen nie.

Hierdie studie dra by tot ons kennis oor die rol wat die genetiese samestelling van die gasheer speel in vatbaarheid vir tuberkulose. Indien die rol van mensgenetika in tuberkulose vatbaarheid korrek verstaan word, kan behandeling van die siekte verbeter word en kan individue wat 'n hoër risiko loop om tuberkulose te ontwikkel geïdentifiseer word.

Acknowledgements

To know even one life has breathed easier because you have lived. This is to have succeeded.

Ralph Waldo Emerson

I would like to thank the following people and institutions for their contribution to this work:

My promoter Prof. Eileen Hoal van Helden and co-promotor Dr. Marlo Möller. Thank you for bestowing your knowledge on a budding young scientist and for being my guiding light, on the dark ocean that is research. Thank you for the encouragement, support, and faith through good and bad. I have learned invaluable scientific lessons from you and would truly not have been able to complete my task without your help. To Marlo, your little baby is finally growing up! I would especially like to thank you for always being there for me and enduring the numerous hardships that I experienced in my research with me. Your support and kindness will never be forgotten.

The head of the department, Prof. Paul van Helden. Thank you for your unequivocal support, both academically and financially.

To my MSc partner, Chandré Wagman. Thank you for all the help, especially the calculations. We may have only known each other for two years, but somehow it feels longer. Thanks also for listening to all my moaning and ranting, especially the last few weeks. It has been tough, but I guess no one would understand it better than you would. Just a few more weeks and then you will be done as well!

To Lance A. Lucas and Khutsokido G. Phalane, my two other lab partners. Thank you for all your support and camaraderie. To Maria Esterhuysen, I still blame you for turning my whole life plan upside down. However, I am grateful for that conversation that we had...COMFORT isn't necessarily good! Also, thank you for helping me with the methylation experiments. To the rest of the Host Genetics Laboratory, thanks for everything.

May Newton-Foot, Michelle Smit for assistance with *E.coli* transformations and culturing and Carrie Kirsten for helping me with the qPCR. Thank you for lending me your expertise and use of your protocols and reagents.

The Division of Molecular Biology and Human Genetics. Thank you all for your friendliness and willingness to help. I will treasure the many friends that I have made.

Our collaborators at the Institute for Clinical Molecular Biology in Kiel, Germany. A special thank you to Prof. Almut Nebel and Dr. Sylvia Hofmann for their assistance in experimental design, as well as manuscript writing.

The Harry Crossley foundation, the National Research Foundation, the South African Medical Research Council and Stellenbosch University for financial support during this project.

The individuals who consented to take part in this research.

Finally to my parents, thank you for all the support and encouragement that you have given me over the years. For always pushing me to be the best that I could be. Thanks for allowing me to follow my dreams and carve my own future. I will always be thankful for all the sacrifices, both financially and personally, and I will be indebted to you for the rest of my life. Shukran!

***This thesis is dedicated to my
parents, Yusuf and Rugaya Salie***

Table of Contents

List of Abbreviations	i
List of Figures	iii
List of Tables	iv
List of Addendums	vi
Chapter 1 – Tuberculosis	
1.1 Brief History of Tuberculosis	1
1.2 Global Epidemic	2
1.3 South African Perspective	2
1.4 <i>Mycobacterium tuberculosis</i> – clinical features and pathogenesis	4
1.4.1 Structure	4
1.4.2 Mode of Transmission	4
1.4.3 Host Immune Response	5
1.4.4 Clinical Manifestation and Diagnosis	5
1.4.5 Treatment	7
1.5 TB – Future Prospects and Fatal Alliances	7
1.5.1 Drug-Resistant TB	7
1.5.2 HIV and TB	8
1.6 Other Risk Factors for the Development of TB	8
1.6.1 Host Genetic Factors	8
1.6.2 Environmental Factors	10
Chapter 2 – Approaches in Disease Gene Identification	
2.1 Current Approaches	12
2.2 Linkage Studies	12
2.3 Association Studies	14
2.3.1 Population-Based Case-Control Association Studies	15
2.3.2 Genome-Wide Association Studies	17
2.3.3 Linkage Disequilibrium and Haplotype Analysis	18
2.4 Animal Models	29
Chapter 3 – Hypothesis and Aims	
3.1 Study Hypothesis	30
3.2 Study Aims	30
Chapter 4 – Methods	
4.1 Reagents and Equipment	31
4.2 Study Participants	31
4.2.1 Study Population	31
4.2.2 Case-Control Samples	31
4.3 DNA samples	32
4.3.1 DNA Extractions	32
4.3.2 Plate Design	32
4.4 Genotyping Methods	33
4.4.1 TaqMan® Genotyping System	33
4.4.2 TaqMan® Protocol	35

4.5 SNP Identification (<i>ANXA11</i>)	35
4.5.1 Exons 4, 5 and 6	37
4.5.2 3' UTR	37
4.5.3 PCR Clean-Up (ExoSAP-IT)	37
4.6 Gene Expression Analysis	37
4.6.1 Sample Selection and Genotyping	37
4.6.2 Cell Culture	37
4.6.2.1 PBMC Isolation	37
4.6.2.2 BCG Infection	38
4.6.3 mRNA Expression	38
4.6.3.1 RNA Extraction	38
4.6.3.2 Reverse-Transcription PCR (RT-PCR)	38
4.6.3.3 Quantitative Real-Time PCR (qPCR)	39
4.6.4 Protein Expression	40
4.6.4.1 Protein Extraction	40
4.6.4.2 Standard Curve	40
4.6.4.3 Western Blot	40
4.7 Methylation Pattern Analysis	41
4.7.1 Sample Selection	41
4.7.2 Bisulfite Conversion of DNA	41
4.7.3 Cloning of Bisulfite Converted DNA	42
4.7.3.1 Amplification and Purification of Bisulfite Converted DNA Fragments	42
4.7.3.2 Ligation to pGEM [®] -T Easy Vector	42
4.7.3.3 <i>E. coli</i> Transformation	43
4.7.3.4 Colony PCR	43
4.7.3.5 Small-Scale Plamid Extraction (Miniprep)	43
4.7.4 Bisulfite Sequencing Analysis	43
4.8 Statistical Analysis	44
4.8.1 Hardy-Weinberg Equilibrium	44
4.8.2 Chi-square Test	44
4.8.3 Fisher's Exact Test	45
4.8.4 Haplotype and Linkage Disequilibrium	45
4.8.5 Power Calculations	45

Chapter 5 – ANXA11

5.1 Annexins	46
5.1.1 Annexin Gene Family	46
5.1.2 Annexin A11	46
5.1.3 Disease Mechanisms – ANXA11	47
5.1.3.1 Cancers	47
5.1.3.2 Sarcoidosis	47
5.2 Results	48
5.2.1 Genotype Analysis	48
5.2.2 Linkage Disequilibrium and Haplotype Analysis	49
5.2.3 Sequencing	50
5.2.4 Gene Expression	51
5.2.5 Methylation	52
5.3 Discussion	53

Chapter 6 – CADM family	
6.1 Cell Adhesion Molecules	58
6.1.1 Cell Adhesion Molecule Family	58
6.1.2 CADM1	58
6.1.3 Other CADM Genes: CADM2, CADM3 and NCAM2	58
6.1.4 Disease Mechanisms – CADM Family	59
6.1.4.1 Cancers	59
6.1.4.2 TB	59
6.2 Results	60
6.2.1 Genotype Analysis	60
6.2.1.1 CADM1	60
6.2.1.2 CADM2	61
6.2.1.3 CADM3	61
6.2.1.4 NCAM2	62
6.2.2 Linkage Disequilibrium and Haplotype Analysis	62
6.2.2.1 CADM1	62
6.2.2.2 CADM3	62
6.3 Discussion	64
Chapter 7 – Conclusions	66
References	68
Appendix	93

List of Abbreviations

5'	Five prime end
3'	Three prime end
X ² test	Chi-square test
μM	Micro molar
AIDS	Acquired immunodeficiency syndrome
AIM2	Absent in melanoma 2
ALG-2	Apoptosis-linked gene-2
ANXA1	Annexin A1
ANXA2	Annexin A2
ANXA7	Annexin A7
ANXA11	Annexin A11
BCG	Bacille Calmette-Guérin
BSA	Bovine serum albumin
C13orf31	Chromosome 13 open reading frame 31
C _q	Threshold cycles
Ca ²⁺	Calcium
CAMs	Cell adhesion molecules
CCDC122	Coiled-coil domain containing 122
CDCV	Common disease, common variant
CRTAM	Class-I-restricted T-cell-associated molecule
DARC	Duffy blood group, chemokine receptor
DC	Dendritic cells
DEPC	Diethyl Pyrocarbonate
df	Degree of freedom
DNA	Deoxyribonucleic acid
GWAS	Genome-wide association studies
H ₀	Null hypothesis
H ₁	Alternative hypothesis
HIV	Human immunodeficiency virus
HIV/AIDS	Human Immunodeficiency Virus/Acquired Immune Deficiency Syndrome
HLA	Human leukocyte antigen
HWE	Hardy-Weinberg equilibrium
IFN-γ	Interferon-γ
Ig	Immunoglobulin
IL-2	Interleukin-2
INH	Isoniazid
kb	Kilo bases
KO	Knock-out
LB	Luria-Bertani
LD	Linkage disequilibrium
MBL	Mannose binding lectin
<i>M. bovis</i>	<i>Mycobacterium bovis</i>
MDR	Multidrug-resistant
MDR TB	Multidrug-resistant tuberculosis
<i>M. tuberculosis</i>	<i>Mycobacterium tuberculosis</i>
NCAM2	Neuronal cell adhesion molecule 2
Necl	Nectin-like

NK	Natural killer
NOD2	Nucleotide-binding oligomerization domain containing 2
PAS	Para-aminosalicylic acid
PBMC	Peripheral blood mononuclear cell
PCR	Polymerase chain reaction
PDZ	Post synaptic density protein (PSD95), <i>Drosophila</i> disc large tumor suppressor (DlgA), and zonula occludens-1 protein (zo-1)
PEM	Protein energy malnutrition
PTST-	Persistently negative tuberculin skin test
qPCR	Quantitative Real-Time PCR
R	Ratio of gene expression change
RB	Reducing Buffer
RIF	Rifampicin
RIP2K	Receptor-interacting serine-threonine kinase 2
RNA	Ribonucleic acid
RPMI-1640	Roswell Park Memorial Institute-1640
RT-PCR	Reverse-Transcription PCR
SAC	South African Coloured
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SLC11A1	Solute carrier 11A member 1
SNP	Single nucleotide polymorphism
T _m	Melting temperature
TB	Tuberculosis
TBM	Tuberculosis meningitis
TBS-T	Tris buffered saline with Tween
TGFβ	Transforming growth factor β
TNF-α	Tumor necrosis factor-α
TNFSM15	Tumor necrosis factor [ligand] superfamily member 15
TST	Tuberculin skin test
UTR	Untranslated region
VDR	Vitamin D receptor
WB	Western Blot
WHO	World Health Organization
WTCCC	Wellcome Trust Case Control Consortium
XDR	Extensively drug resistant
XDR TB	Extensively drug resistant tuberculosis

List of Figures

Figure 1: Estimated TB incidence rates by country, 2007	3
Figure 2: Estimated TB incidence rates in South Africa and its provinces, 1999 – 2006	3
Figure 3: Scanning electron micrograph image of <i>M. tuberculosis</i>	4
Figure 4: Estimated HIV prevalence in new TB cases, 2007	8
Figure 5: Current strategies for the identification of TB susceptibility genes	12
Figure 6: Current understanding of genes involved in altered susceptibility to <i>M. tuberculosis</i> infection	16
Figure 7: Schematic representation of sample allocation in 384 well plates	32
Figure 8: Schematic representation of the TaqMan® Genotyping System Flowthrough	33
Figure 9: Graphical representation of the TaqMan® Genotyping Assay	33
Figure 10: Cluster plot used for the calling of genotypes	34
Figure 11: Plot of LD between the <i>ANXA11</i> markers in the SAC population, generated by Haploview v4.1	49
Figure 12: 1.5% agarose gel with amplified PCR products	50
Figure 13: Alignment of sequence data using the Sequencher v4.7 software	50
Figure 14: Differences in mRNA levels between individuals with different genotypes	51
Figure 15: 2% agarose gel with amplified bisulfite treated PCR products of the promoter region of <i>ANXA11</i>	52
Figure 16: Blue/White colony selection of <i>E. coli</i> JM109 cells	52
Figure 17: Colony PCR of <i>E. coli</i> JM109 cells for the verification of DNA fragment of interest	53
Figure 18: Lollipop diagram, with circles representative of CpGs, indicating no differential methylation between cases and controls	53
Figure 19: Structure of <i>CADM1</i> protein, showing (left to right) the 3 Ig domains; the transmembrane domain and the short cytoplasmic domain	59
Figure 20: Plot of LD between the <i>CADM1</i> markers in the SAC population, generated by Haploview v4.1	63
Figure 21: Plot of LD between the <i>CADM3</i> markers in the SAC population, generated by Haploview v4.1	63

List of Tables

Table 1: Different stages of TB infection	6
Table 2: Diagnostic tests currently employed in TB	6
Table 3: Anti-TB drugs currently in use.	7
Table 4: Twin studies investigating the heritability of TB	9
Table 5: Chromosomal regions identified by genome-wide linkage analysis for susceptibility to TB	14
Table 6: Association studies investigating TB susceptibility candidate genes	19
Table 7: Murine genes which result in increased susceptibility to mycobacteria species when disrupted	29
Table 8: Characteristics of individuals included in the case-control association studies	31
Table 9: TaqMan® Genotyping Assays used in this study	35
Table 10: Primer sets and their respective sizes used to sequence <i>ANXA11</i> regions	36
Table 11: Primer set used to sequence the rs7071579 polymorphism	37
Table 12: Primer sets used for qPCR analysis	39
Table 13: Primers used for amplification of CpG island in promoter region of <i>ANXA11</i>	42
Table 14: <i>ANXA11</i> polymorphisms investigated as TB susceptibility variants in the SAC population	48
Table 15: Statistical analysis of <i>ANXA11</i> polymorphisms	48
Table 16: Haplotype analysis for <i>ANXA11</i> polymorphisms	49
Table 17: Polymorphisms identified during sequencing of <i>ANXA11</i> gene regions	51
Table 18: Comparison of the allele frequencies in the German, South African Coloured and Yoruba control populations	57
Table 19: Various names used in the literature for the <i>CADM</i> genes	58
Table 20: <i>CADM</i> polymorphisms investigated as TB susceptibility variants in the SAC population	60
Table 21: Statistical analysis of <i>CADM1</i> polymorphisms	60
Table 22: Statistical analysis of <i>CADM2</i> polymorphisms	61
Table 23: Statistical analysis of <i>CADM3</i> polymorphisms	61
Table 24: Statistical analysis of the <i>CADM3</i> polymorphism, rs12057331, and susceptibility to TBM	62
Table 25: Statistical analysis of <i>NCAM2</i> polymorphisms	62

List of Addenda

Addendum 1: Buffers, Solutions and Gels	93
Addendum 2: Reagents	96
Addendum 3: Equipment	99
Addendum 4: Software Packages	101
Addendum 5: Online Resources	101

Chapter 1: Tuberculosis

The capacity to blunder slightly is the real marvel of DNA. Without this special attribute, we would still be anaerobic bacteria and there would be no music.

Lewis Thomas

1.1 A Brief History of Tuberculosis

Tuberculosis (TB) has recently re-emerged as a major global health concern. However, TB has been plaguing humanity for centuries before this, with the earliest documentation of TB occurring in Egypt as early as 5000 years ago, based on the isolation of *Mycobacterium tuberculosis* (*M. tuberculosis*) DNA from mummies^{1, 2}. Earlier names for TB included phthisis, which means consumption or to waste away, and was identified as the most rife disease of the times by Hippocrates³. He also noted that the disease occurred more frequently in individuals between the ages of 18 and 25 years old, and almost always resulted in death. It was Clarissimus Galen, a Greek physician, who described phthisis as an “ulceration of the lungs, chest or throat, accompanied by coughs, low fever, and wasting away of the body because of pus”; and described it as a disease of malnutrition⁴.

With the commencement of the 17th century, Europe was struck by a TB epidemic which lasted for 200 years, and was later known as the “Great White Plague”³. It is believed that overcrowding and poor sanitary conditions that were characteristic of the rapidly growing cities of the Western World provided the necessary milieu for the spread of this airborne pathogen. Due to the exploration and colonization that was typical of this period, the TB epidemic slowly consumed the colonized nations as well⁵. Although it is believed that TB existed in America and Africa before the arrival of the Europeans, the disease was very rare among the indigenous people, but after contact with the Europeans, the mortality rate due to TB rapidly increased within these native populations.

Even though 17th century Europe was plagued with TB disease, it is the century during which scientists began to unravel the mysteries of the disease and its causative agent^{1, 5, 6}. It was Franciscus Sylvius de la Bøe of Amsterdam who was the first to identify the presence of tubercles in the lungs of TB patients as a characteristic of the disease, and this finding was later corroborated by the English physician Richard Morton³. Both de la Bøe and Morton also believed that the disease was hereditary; however, Morton also considered transmission by intimate contact as a possible mechanism. Later it was shown by Gaspard Laurent Bayle that the tubercles noted by de la Bøe and Morton were not the products of the disease but the actual cause, which gave rise to the name of the disease used today: tuberculosis.

During the 18th and 19th centuries, various physicians and epidemiologists made scientific breakthroughs in that they finally discovered that the causative agent of TB was a microorganism, and that it could be transmitted between individuals and between humans and various other mammal species. In 1882, Robert Koch made his famous presentation, in which he showed that the bacterium, *M. tuberculosis*, was the causative agent of tuberculosis disease^{2, 3}. Towards the end of the 19th century, Koch announced the isolation of a compound that inhibited the growth of the tubercle bacilli when administered to guinea pigs both pre- and post-exposure. This compound was called ‘tuberculin’, and soon after its discovery it was used as a therapeutic vaccine in a clinical trial, the results of which were extremely disappointing. However, it was found to be valuable in the diagnosis of TB, and later gave rise to the currently used Tuberculin Skin Test/Mantoux test, which is used for diagnosis of latent TB.

The 20th century brought with it the advent of the first successful vaccine and the use of chemotherapy in the fight against TB. At the beginning of the century, Albert Calmette and Camille Guérin successfully created the attenuated strain of *M. bovis*, Bacille Calmette-Guérin (BCG), which was avirulent in cattle, horses, rabbits and guinea pigs¹. In 1943, the first antibiotic,

streptomycin, was isolated by Selman A. Waksman and Albert Shatz from *Streptomyces griseus*. The drug was effective against *M. tuberculosis in vitro*, and post-infection in guinea pigs. In 1944, the first human was treated with the drug and two clinical studies followed, one in Europe and the other in the United States of America. It was noted that individuals who were treated with streptomycin exhibited a substantial improvement in the outcome of disease. The findings of these studies were however of a “double-edge sword” nature, in that the investigators also noted that following the first months of treatment, some patients’ disease status began to worsen; from which they concluded that the pathogen was able to develop resistance to the drug ⁷. This was soon followed by the administration of para-aminosalicylic acid (PAS) as an oral therapy ³. PAS therapy was successful in treating TB, and unlike streptomycin, it was non-toxic and the bacterium was not able to develop resistance to the drug easily. In the late 40’s, researchers noted that when TB patients were treated with both streptomycin and PAS, the disease outcome was much more favourable than when patients were treated with only one of the drugs. The 1950’s saw the introduction of isoniazid (INH) as a chemotherapeutic agent against TB. Unfortunately, however, like its predecessor streptomycin, it was found that *M. tuberculosis* was able to easily develop resistance to this drug. On the positive side, treatment with streptomycin, PAS and INH was found to be highly effective in the treatment of TB, and for the first time TB was curable. This led to a wave of new drugs being developed for the treatment of TB, including rifampicin, pyrazinamide, ethambutol, cycloserine and ethionamide; all of which are still in use today for the treatment of TB. However, the problem of drug resistance acquisition continues.

TB has had a long history, claiming millions of victims during the ages and taking more lives than any other microbial disease. Due to this, TB has left its mark on humanity; in music, art and literature and playing a major role in the advancement of biomedical sciences and healthcare. So what does the 21st century hold for this highly efficient pathogen?

1.2 Global Epidemic

In 1993, the World Health Organization (WHO) declared TB a global health emergency ⁸. Approximately one-third of the world’s population is infected with *M. tuberculosis*, with the WHO estimating that 9.27 million people developed TB in 2007, with around 2 million deaths. Most TB incident cases were from developing countries, predominantly Asia and Africa (Figure 1) with India, China, Indonesia, Nigeria and South Africa rated as the five countries with the highest TB burden respectively.

TB is the world’s second most common cause of death by an infectious agent, after the human immunodeficiency virus/acquired immune deficiency syndrome (HIV/AIDS). Although TB is “under control” in developed nations, it remains a major health threat in third world countries. This is predominantly due to high HIV infection rates, poor health care, poor socio-economic status and the development of drug resistant strains of *M. tuberculosis*.

1.3 South African Perspective

TB has been a long-standing health issue in South Africa and with the current HIV epidemic sweeping the nation, the fight for the eradication of TB has become even more difficult, with a HIV prevalence rate of 18.1% (2007) in South African adults ¹⁰. As mentioned previously, South Africa is currently ranked 5th in the world with regards to TB burden, with TB treatment success rates remaining low due to the high number of deaths due to TB, an increase in relapses due to poor adherence to treatment therapy and the spread of multidrug-resistant (MDR) and extensively drug

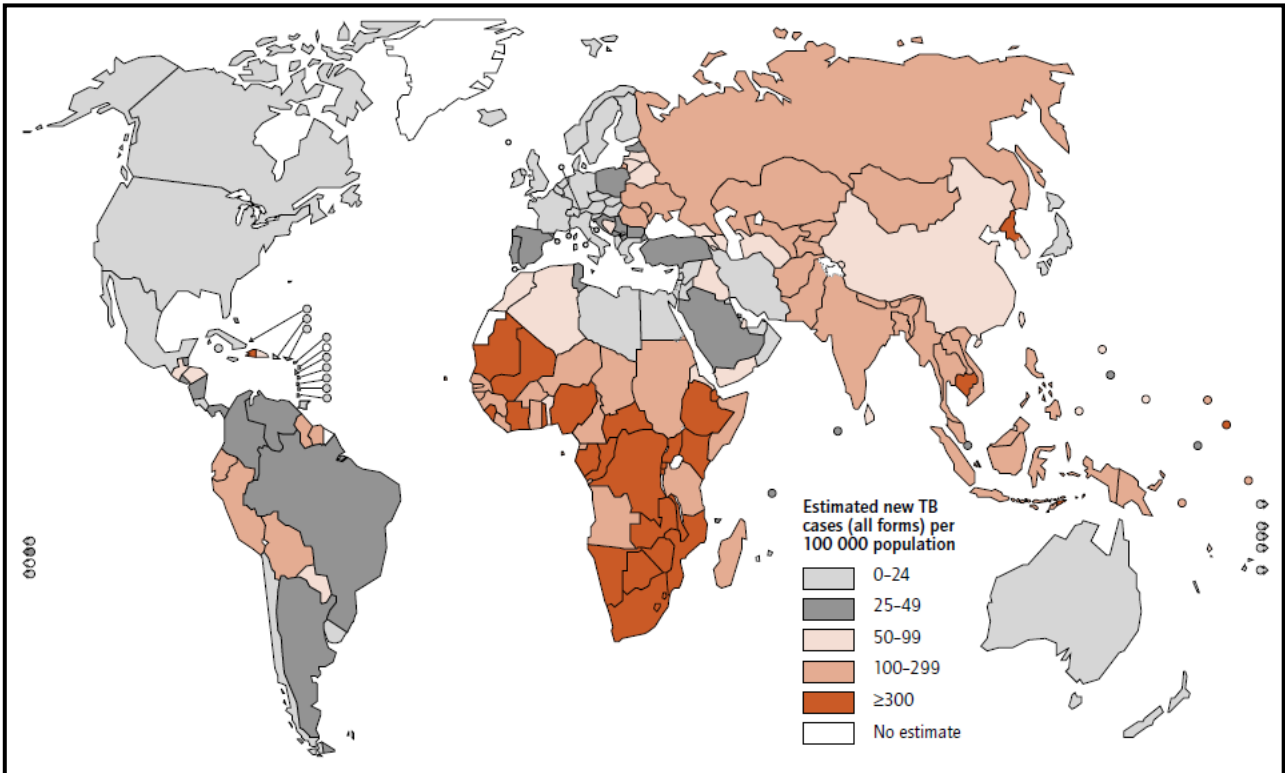


Figure 1: Estimated TB incidence rates by country, 2007 (Reproduced ⁸)

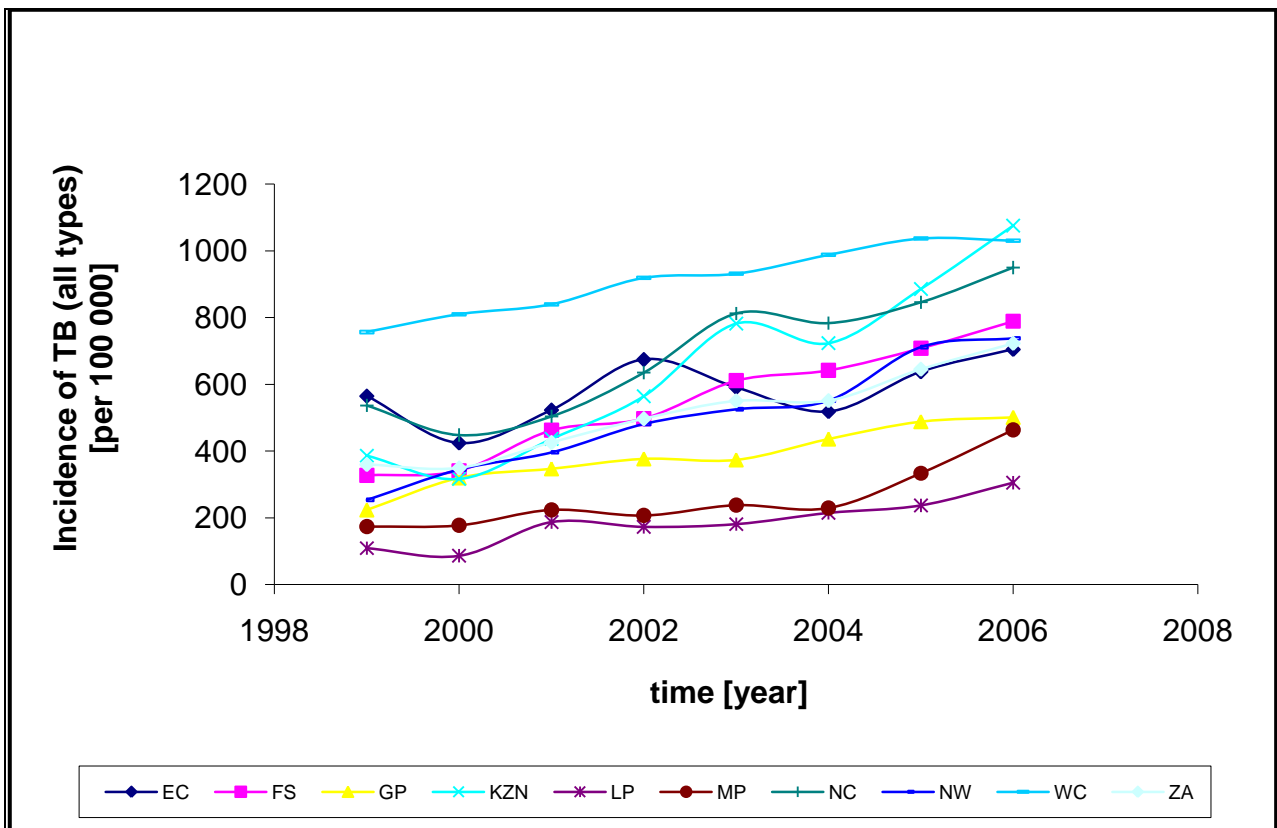


Figure 2: Estimated TB incidence rates in South Africa and its provinces, 1999 – 2006 ⁹

resistant (XDR) TB ¹¹.

Of the South African provinces, the Western Cape has maintained a consistently high TB incidence rate 1 030 per 100 000, only surpassed in 2006 by the Kwa-Zulu Natal Province with 1 076 per 100 000. The current incidence rate for TB in South Africa is 948 per 100 000 individuals ⁹ (Figure 2).

1.4 Mycobacterium tuberculosis – clinical features and pathogenesis

1.4.1 Structure

TB is a result of infection by the pathogen, *M. tuberculosis*, although other members of the *Mycobacterium tuberculosis* complex are also known to result in TB, including *M. africanum*, *M. bovis* and *M. microti* ^{12, 13}. These bacteria are rod-shaped (Figure 3), non-spore-forming, aerobic, Gram-positive bacteria. Mycobacteria usually measure 0.5 µm x 0.3µm and are classified as acid-fast bacilli due to the difficulty with which dyes are removed from the cell wall with the treatment of acid-alcohol after the staining of the cell wall ^{14, 15}.



Figure 3: Scanning electron micrograph image of *M. tuberculosis*. Courtesy of the Centre for Disease Control.

The cell wall structure of *M. tuberculosis* is essential for the survival of this pathogen intracellularly ¹⁴. The cell wall of *M. tuberculosis* is comprised of mycolic acid (fatty acid) which is covalently bound to arabinogalactan (peptidoglycan-bound polysaccharide) which gives rise to the bacterium's extraordinary lipid barrier. This lipid barrier is largely responsible for the ability of *M. tuberculosis* to develop resistance to antibiotics and evade the host's defence mechanisms. The presence of lipoarabinomannan on the cell wall of the bacterium confers upon it its immunogenic properties, allowing the bacterium to survive within macrophages, while the composition and quantity of the bacterium's cell wall components directly influences its pathogenicity and growth rate.

1.4.2 Mode of Transmission

M. tuberculosis is an airborne pathogen with transmission as the result of the spread of small airborne droplets, droplet nuclei, usually generated by coughing, sneezing or talking of an infected person with pulmonary tuberculosis ². Due to the small size of these droplets, they are able to remain airborne for long time periods. When these droplet nuclei enter the lungs of an uninfected individual various outcomes are possible. Depending on the pathogenicity of the *M. tuberculosis* strain and the host's immune response ¹⁶, the newly infected individual can go on to develop

active TB disease, prevent the growth and spread of the bacteria or immediately kill the bacteria. Various factors can influence the transmission of *M. tuberculosis*, including the number of bacilli contained within the droplet nuclei, virulence of the bacilli and ventilation.

1.4.3 Host Immune Response

Cell-mediated immunity is the primary response employed by the host to fight off *M. tuberculosis* infection¹². TB infection starts when the bacteria reach the alveoli and are phagocytosed by alveolar macrophages or dendritic cells¹⁷. Depending on the virulence of the infecting strain and the immune system of the host there are two possible outcomes^{15, 17, 18}. Firstly, as occurs in most individuals, phagocytosis of the bacteria by macrophages results in the initiation of a strong host immune response and the subsequent death of the mycobacteria. However, if the host is unable to contain the growth of the bacteria the formation of granulomas will follow. They have the effect of limiting the growth and further spread of the bacilli and are formed by macrophages, T lymphocytes, B-lymphocytes and fibroblasts¹⁷. The T lymphocytes, which surround the infected macrophages, release various cytokines including interferon gamma (IFN- γ) which activates the macrophages to destroy the bacterium. In most cases however, the bacteria are not killed but become dormant resulting in latent infection^{17, 19}. At this stage of infection, the immune system effectively contains the infection. However, if it is unable to, the bacteria will begin to actively replicate, resulting in necrosis of the infected lung tissue and further spread of the bacteria to other body organs (extrapulmonary TB) or to new hosts.

1.4.4 Clinical Manifestation and Diagnosis

Based on the immune response that is elicited at the point of infection, TB disease can manifest in various forms; including latent infection, primary disease, active TB disease and extrapulmonary TB disease (Table 1). Each stage of disease can be characterised by its own set of symptoms and means of diagnosis¹⁴.

In summary, latent TB occurs when the host immune system is unable to completely eliminate the bacterium after infection but is able to contain its growth within an enclosed location. Although these bacteria are viable and are able to persist for many years, they do not result in active TB disease and therefore no symptoms are experienced and these individuals are not infectious. In the past, latent TB was diagnosed by means of the tuberculin skin test (TST), however, due to false-negative (immunocompromised or malnourished individuals) and false-positive (response to BCG vaccination) skin tests, this method has been replaced by the QuantiFERON-TB Gold test which is thought to be more sensitive and time efficient (Table 2).

Active TB disease occurs when the immune response initiated by the host is unable to control the infection, resulting in the active growth of the bacterium in the lungs of the infected individual. These individuals experience various symptoms which are indicative of active TB disease. These include fatigue, weight loss, fever, extensive coughing, night sweats and anaemia. These individuals are highly infectious. Diagnosis of active TB usually involves chest X-rays in addition to sputum smears and sputum cultures (the "gold standard"). However, recent advances in molecular biology have allowed for the development of faster diagnostic tests. Amplification of DNA and RNA now allows for rapid detection of microorganisms (Table 2).

Table 1: Different stages of TB infection

Early Infection	Early Primary Progressive (active)	Late Primary Progressive (active)	Latent
Immune system fights infection	Immune system does not control initial infection	Cough becomes productive	Mycobacteria persist in the body
Infection generally proceeds without signs or symptoms	Inflammation of tissue ensues	More signs and symptoms as disease progresses	No signs or symptoms occur
Patients may have fever, paratracheal lymphadenopathy, or dyspnea	Patients often have nonspecific signs or symptoms (e.g., fatigue, weight loss, fever)	Patients experience progressive weight loss, rales, anaemia	Patients do not feel sick
Infection may be only subclinical and may not advance to active disease	Non-productive cough develops	Findings on chest radiograph are normal	Patients are susceptible to reactivation of disease
	Diagnosis can be difficult: findings on chest radiographs may be normal and sputum smears may be negative for mycobacteria	Diagnosis is via cultures of sputum	Granulomatous lesions calcify and become fibrotic, become apparent on chest radiographs
			Infection can reappear when immunosuppression occurs

Finally, extrapulmonary TB occurs when the infection can no longer be contained within the lungs and the bacterium enters the blood system (miliary TB) and is able to infect other organs. The most seriously infected region is the central nervous system, which results in tuberculosis meningitis (TBM). This form of TB is often fatal and is usually characterized by headaches and mental instability. Miliary TB on the other hand is much more difficult to diagnose, due to the nonspecific symptoms experienced by these individuals; including fever, weight loss and weakness.

Table 2: Diagnostic tests currently employed in TB

Variable	Sputum smear	Sputum culture	Polymerase chain reaction (PCR)	Tuberculin skin test	QuantiFERON-TB test	Chest radiography
Purpose of test study	Detect acid-fast bacilli	Identify <i>M. tb</i>	Identify <i>M. tb</i>	Detect exposure to mycobacteria	Measure immune reactivity to <i>M. tb</i>	Visualize lobar infiltrates with cavitation
Time required for results	<24 hours	3-6 weeks with solid media, 4-14 days with high-pressure liquid chromatography	Hours	48-72 hours	12-24 hours	Minutes

1.4.5 Treatment

The advent of sanatoria represented the first widely used treatment for the fight against TB^{1, 3}. Treatment was rather simple, with infected individuals receiving good nutrition and maximum exposure to fresh air. However, with the discovery of the antibiotics, streptomycin and PAS in the 1940's, current TB treatment has altered drastically³. Today, treatment goals include curing infected individuals and limiting the chance of relapse, stopping transmission and preventing the development of drug resistance and death²⁰.

Current TB drug treatments require extended time periods (usually six months) of medical drug use²⁰. TB treatments are usually composed of 2 phases, namely an initial phase (2 months) and a continuation phase (4-7 months). During the initial phase infected individuals receive four first-line* drugs (Table 3); while during the continuation phase these individuals receive only isoniazid (INH) and rifampicin (RIF). However, due to this long treatment plan, poor treatment adherence is observed which leads to the development of drug-resistant TB. In this case, patients are treated with second-line drugs (Table 3) for an extended time period.

Table 3: Anti-TB drugs currently in use.

First-line drugs	Second-line drugs
Isoniazid*	Cycloserine
Rifampicin*	Ethionamide
Ethambutol*	Levofloxacin
Pyrazinamide*	Moxifloxacin
	Gatifloxacin
Rifapentine	p-Aminosalicylic acid
Rifabutin	Streptomycin
	Amikacin/Kanamycin
	Capreomycin

*First-line drugs by default

Although TB treatment plans are broadly applicable, treatment modifications should be made under certain conditions; including HIV infection, drug-resistance, pregnancy and the treatment of children.

1.5 TB – Future Prospects and Fatal Alliances

1.5.1 Drug-Resistant TB

The development of resistance in *M. tuberculosis* to various anti-TB drugs is considered one of the draw-backs of the use of chemotherapy to fight TB disease²¹. Drug resistance is defined as the inability of otherwise effective drugs to kill the bacterium and is a result of drug misuse and mismanagement. This includes poor treatment adherence, incorrect prescriptions by health-care workers (wrong treatment, dosage, and treatment period), the unavailability of drugs and drugs of poor quality.

There are currently two forms of drug-resistant TB, namely multidrug-resistant TB (MDR TB) and extensively drug resistant TB (XDR TB)^{22, 23}. MDR TB is defined as TB that is resistant to at least two of the most efficient first line drugs, INH and RIF. These TB cases are usually much more difficult to treat, with a mortality rate of approximately 40% - 60%. XDR TB on the other hand is TB that is resistant to both INH and RIF, in addition to any fluoroquinolone and at least one of the

three injectable second-line drugs (amikacin, capreomycin or kanamycin)²⁰. This form of TB is much harder to treat, mainly due to the resistance of the pathogen to most first- and second-line anti-TB drugs, resulting in poorly effective treatment options and outcomes.

1.5.2 HIV and TB

With the global spread of the HIV/AIDS epidemic, controlling TB disease has become exceedingly difficult. Due to the immunocompromising effect of HIV infection, TB disease prognosis has worsened, with noted increasing risks for reactivation of latent *M. tuberculosis* infection and rapid disease progression^{24, 25}. It has also been noted that individuals only infected with *M. tuberculosis* exhibit lifetime risks of developing TB of between 10% and 20%, while individuals infected with both *M. tuberculosis* and HIV have an annual risk of developing TB greater than 10%.

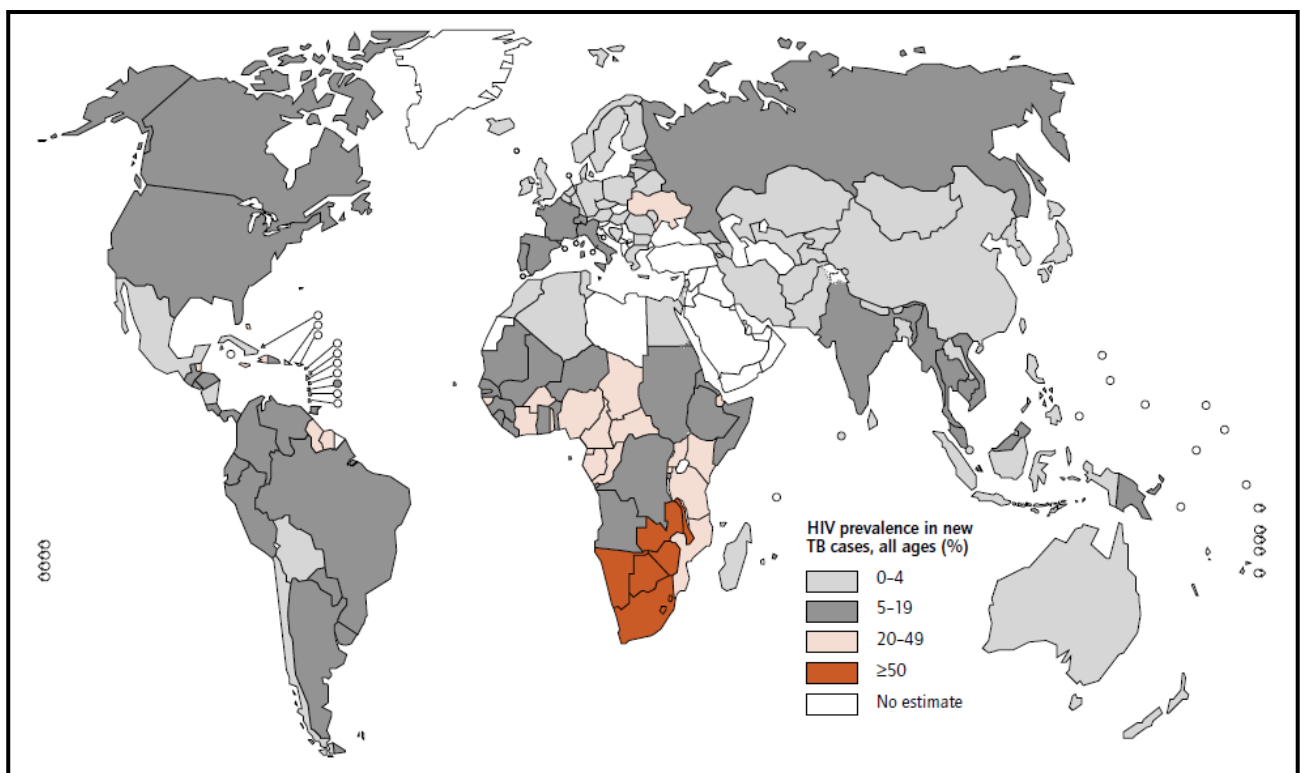


Figure 4: Estimated HIV prevalence in new TB cases, 2007⁸

Southern Africa currently has the highest prevalence of HIV infection in new TB cases (Figure 4), with more than 50% of new TB cases being co-infected with both organisms¹¹.

1.6 Other Risk Factors for the Development of TB

1.6.1 Host Genetic Factors

There is clear evidence that host genetic factors play a crucial role in the development of active clinical tuberculosis and that pathogenic factors, although important, is not the sole deciding factor in who progresses to disease and who does not. Current global TB statistics highlight the important fact that of the one-third of the world's population infected with TB; only 10% will go on to develop active disease. This clearly shows that host genetic factors are extremely important in the outcome of this infection.

To support this idea, various classical studies have been conducted on the effects of host genetics in infectious diseases; these include numerous twin and adoption studies. Twin studies have been used to determine how susceptibility to infectious diseases differs between monozygotic and dizygotic twins²⁶⁻²⁸. All of these studies found that monozygotic twins, who share the exact genetic make-up, have a higher concordance for disease than dizygotic twins, whose genetic make-up differs from each other (Table 4). This illustrates that host genetic factors are major contributors to the development of TB since twins generally share the same environment. Adoption studies have also been conducted, which have conclusively shown that adopted children are more likely to die from an infectious disease if their biological parents died from an infectious disease compared to their adoptive parents, again highlighting the importance of host factors over environment in the outcome of disease²⁹⁻³².

Table 4: Twin studies investigating the heritability of TB.

Study	Number of Twins		% Concordance		Reference
	Monozygotes	Dizygotes	Monozygotes	Dizygotes	
Diehl <i>et al</i> , 1936	80	125	65	25	33
Uehlinger <i>et al</i> , 1938	12	34	58	6	34
Kallmann <i>et al</i> , 1943	78	230	62	18	27
Harvald <i>et al</i> , 1965	135	513	37	15	35
Comstock <i>et al</i> , 1978	54	148	32	14	26
Simonds, 2004	55	150	32	14	28
van der Eijik, 2007	54	148	21	19	36

Other evidence supporting the idea that host genetics plays an important role in susceptibility to TB, include the rather unfortunate event which occurred in Lübeck, Germany in 1926, where 251 newly born infants were mistakenly vaccinated with a live virulent strain of *M. tuberculosis* rather than an attenuated strain³⁷. Of the 251 babies; 47 went on to develop latent TB disease, 77 died from TB, and 127 had radiological signs of TB but later recovered. This incident demonstrates that certain individuals in a population have an efficient innate immune response to TB. Another event which gives credence to the role of host genetics in susceptibility to TB comes from the initial exposure of a population to *M. tuberculosis*. This was particularly seen in the Qu'Appelle Indians who, when first exposed to the bacterium, had a high annual TB mortality rate (10%), but after 40 years, the annual death rate due to TB dropped significantly (0.2%)³⁸. This incident can be interpreted as illustrating the effects of strong selection pressures against those genes that confer susceptibility to TB, and that the causative genetic variants within these genes are selected against and not transmitted to subsequent generations, resulting in less susceptible future populations. This effect of natural selection can also be seen in other populations. In the case of Europeans and Africans, it appears that individuals from Europe are less susceptible to TB infections, whereas individuals from Africa seem to be more prone to infection by *M. tuberculosis*. It is believed that this is due to European populations having been exposed to *M. tuberculosis* for centuries (White Plague, 17th century), resulting in a more resistant population³¹. In Africa on the other hand, exposure to *M. tuberculosis* occurred rather recently and with the availability of drugs to fight *M. tuberculosis* infection, natural selection has not been able to remove susceptibility genetic variants from the population. Although environmental and socio-economic factors differ immensely between these two populations, they alone cannot account for this population variation. This has been shown in a study conducted in a USA nursing home which found that individuals of African ancestry were twice as likely to be infected with *M. tuberculosis* compared to individuals of European ancestry, even though they shared the same environment³⁹.

All these studies and historical incidents clearly demonstrate that human genetic variation plays a key role in the disease outcomes.

1.6.2 Environmental Factors

Environmental factors also play an important role in the outcome of TB disease, and these factors include socio-economic, nutrition, smoking and alcohol abuse.

Various studies have showed that disadvantaged communities tend to have a higher incidence of TB, which is mainly attributable to poverty and the associated overcrowded living conditions^{40, 41}. This can be seen in a report (1995) which listed “high risk environments” for TB, in which prisons, nursing homes and homeless shelters were included. The role of overcrowded living conditions in TB development was also highlighted when it was noted that there was a higher incidence of TB in monasteries and refugee camps, which became overcrowded due to the flight of refugees during the Chinese occupation of Tibet. This has also been seen in America, where racial segregation resulted in over-crowding and limited health care access in minority areas. A study done by Farmer (1997) also showed that treatment compliance was essentially determined by economic factors⁴².

As mentioned previously, nutrition is another environmental factor that plays an important role in disease development. In 2004, Cegielski *et al.* conclusively showed for the first time the relationship between malnutrition and TB, based on studies in humans and experimental animals. Malnutrition may alter cell-mediated immunity, which is the principle host defence mechanism against TB⁴³. In the guinea pig TB model, various studies have also shown that chronic protein energy malnutrition (PEM) has a negative effect on the immunity to *M. tuberculosis*. PEM results in significantly reduced lymphocyte stimulation, in addition to low level secretion of the Th1 cytokines IL-2, IFN- γ and TNF- α . Additionally, it was noted that PEM animals generated macrophages which produced higher levels of transforming growth factor β (TGF β), which results in the suppression of T cells and inflammation⁴⁴. With regards to vegetarianism and the risk of developing TB, Finch *et al.*, (1991) showed that in a retrospective study of TB in an Indian subcontinent population, Hindu Asians were at a greater risk for developing TB when compared to Asian Muslims with an overall incidence ratio of 4.5⁴⁵. Strachan *et al.*, (1995) using a case-control study technique in the same population, showed that this increased risk for developing TB was in fact due to diet and not religion, since vegetarianism was common practice in Hindus but not Muslims⁴⁶. After adjusting for diet (vegetarianism) other factors such as socioeconomic status, migration, lifestyle choices, age and sex made little difference to the relative risk of developing TB in this population. They also showed that individuals who were lactovegetarians had an 8.5 fold risk of developing TB when compared with daily meat/fish eaters. From this they concluded that a vegetarian diet is an independent risk factor for developing TB, and postulated that it could be due to impairment of the immune system through the deficiency of micronutrients. One of these micronutrients is vitamin D^{47, 48}, which has been shown to have an immunoregulatory role in both lymphocytes and monocytes and a deficiency in vitamin D could lead to an impaired host defence to *M. tuberculosis*⁴⁹.

Both smoking and alcohol abuse contribute to the development of TB disease⁵⁰⁻⁵³. It has also been shown that children who live with adults who previously had TB and were exposed to second hand cigarette smoke were at a higher risk of developing TB⁵³. Interestingly, in India, it was found that individuals with TB were three times more likely to be smokers when compared to the rest of the population. Smokers tend to have faster TB disease progression, poor treatment adherence and are more likely to relapse. For alcohol abuse, it has been noted that such individuals have a

relative risk of 3 (95% CI: 1.89-4.59) with regards to developing TB⁵¹. This is believed to be due to the pathogenic impact of alcohol on the immune system. Alcohol abuse can alter the pharmacokinetics of the medication used in the treatment of TB, as well as resulting in higher rates of re-infection and treatment defaults which also increases the risk of developing drug-resistant TB. Analysis of current data also indicates that approximately 10% of the global TB cases are attributable to heavy alcohol consumption⁵⁰⁻⁵².

These risk factors indicate the complexity of TB disease, and show that although *M. tuberculosis* is necessary, it is not sufficient for the development of clinical TB disease. It also highlights the necessity of integration of epidemiology, host genetics and environmental factors if we are to successfully eradicate this disease.

Chapter 2: Approaches in Disease Gene Identification

Science is facts; just as houses are made of stones, so is science made of facts; but a pile of stones is not a house and a collection of facts is not necessarily science.

Henri Poincaré

2.1 Current Approaches

There are two chief approaches currently being used in the identification of susceptibility genes for human TB, namely population-based gene association studies and family-based linkage analysis (Figure 5) ³². However, both study designs have advantages and disadvantages, with the use of both methodologies combined being most likely to yield success. Identification of genes involved in complex diseases can either be based on a hypothesis (association studies) or not (linkage studies) ⁵⁴.

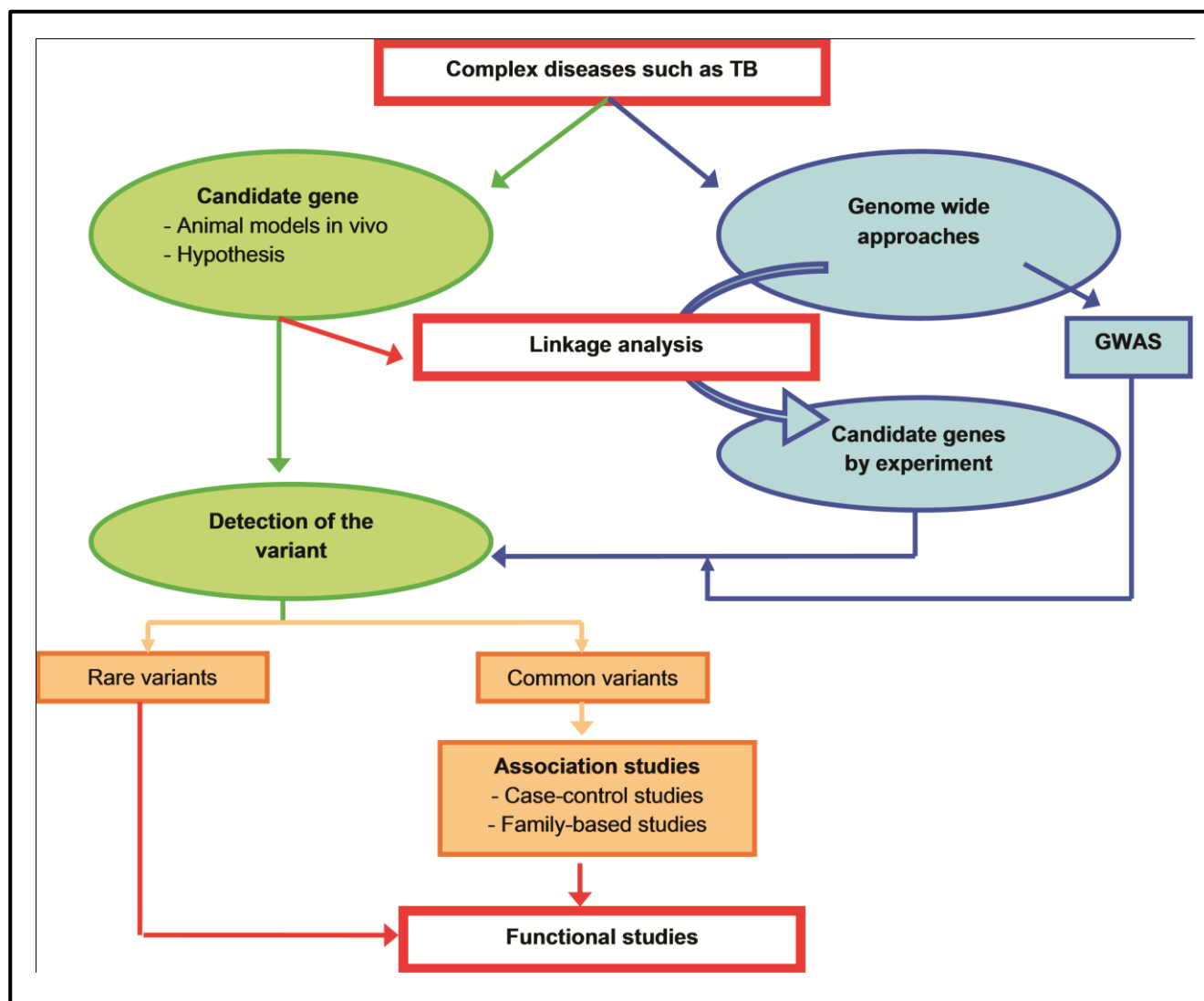


Figure 5: Current strategies for the identification of TB susceptibility genes ³²

However, recent advances in genotyping techniques and molecular biology have resulted in the introduction of novel techniques for the identification of susceptibility genes for TB. These include genome-wide association studies (GWAS) ⁵⁵, the use of admixture mapping ³² and epigenetic studies ⁵⁶.

2.2. Linkage Studies

Linkage studies are used to identify chromosomal regions that contain TB susceptibility genes by testing for co-segregation between a genetic marker and a possible disease locus. This method of gene identification requires a large number of families with affected children ³⁷. The advantage of

this method lies in that it can evaluate the entire genome or focus on a specific region in the genome. The former allows for the identification of novel genes and pathways that would previously not have been considered. Linkage studies are based on the assumption that chromosomal regions and the disease of interest segregate non-randomly, allowing for the identification of these regions in large affected families. Once linkage has been identified in a region, further studies are conducted to narrow down the interval on the chromosome so that the gene can be identified, possibly by positional cloning⁵⁷. Recent advances have allowed researchers to employ fine-mapping for the identification of the susceptibility gene. Fine-mapping involves the use of gene-associated markers (single nucleotide polymorphisms, SNPs) and whether or not they are transmitted with the disease in affected offspring⁵⁸. There are however disadvantages to this method³⁷. Firstly, genome-wide linkage analysis requires large number of families with at least two affected children, which is not easily attainable. This method also has lower statistical power compared to association studies, mainly due to the difficulty of attaining multi-case families than random cases, and is therefore likely to identify regions containing genes with modest effects. This method is also better suited for gene identification of monogenic diseases in that it is able to identify a single chromosomal region which can be narrowed down to identify the causative gene³². Complex diseases such as TB on the other hand, which involves numerous genes, may not be best suited for gene identification using this method.

To date, seven genome-wide linkage studies have been conducted to identify genes associated with varied susceptibility to TB (Table 5). The first genome-wide linkage scan was conducted in the Gambian and South African populations using sib pairs⁵⁹. This study identified the Xq and 15q chromosomal regions in the respective populations. Fine-mapping of chromosome 15q11-13 region identified the ubiquitin protein ligase E3A (*UBE3A*) or another closely linked gene as a possible susceptibility gene for TB⁵⁸. The identification of the Xq chromosomal region in Gambians as a region containing a possible susceptibility gene for TB is also interesting, as current statistical data shows that males have a higher incidence of developing TB⁶⁰. This effect however could be attributable to other non-genetic factors as well⁵⁹. To date, one gene on the X chromosome has been found to be associated with susceptibility to TB in the general population, namely toll-like receptor 8 (*TLR8* in Indonesia and Russia)⁶¹. A genome-wide linkage scan conducted in the Ugandan population identified four chromosomal regions, of which two were found to contain regions, 2q21-q24 and 5p13-5q22, associated with a persistently negative tuberculin skin test (PTST-)⁶². Interestingly, the most recent genome-wide linkage scan conducted in the South African population identified two chromosomal regions, 5p15 and 11p14, to be associated with various tuberculin skin test (TST) properties⁶³. The 11p14 region was found to be involved in controlling human resistance to *M. tuberculosis* infection. On the other hand, the 5p15 region was found to be involved in determining the extent of the TST, which supports the findings of the Ugandan genome-wide linkage scan that identified the 5p13-5q22 region to be involved in PTST-⁶². Fine mapping of this 5p15 region resulted in the identification of the solute carrier family 6, member 3 (*SLC6A3*) gene as a potential candidate for the regulation of TST intensity⁶³.

Of the seven genome-wide linkage scans conducted in TB, very little overlap between the identified susceptibility regions has been observed⁶⁴. This could be due to differences in study designs, for example, low sample numbers, differences in phenotype/diagnostic criteria or population specificity.

Table 5: Chromosomal regions identified by genome-wide linkage analysis for susceptibility to TB

Population	Chromosomal region	TB* phenotype	Reference
South Africa	15q11-q13	TB	59
The Gambia	Xq	TB	
Brazil	10q26.13	TB	65
	11q12.3	TB	
	20p12.1	TB	
Morocco	8q12-q13	TB	66
South Africa	6p21-q23	TB	67
Malawi	20q13.31-33	TB	
Uganda	2q21-q24	PTST-*	62
	5p13-5q22	PTST-	
	7p22-p21	TB	
	20q13	TB	
Thailand	5q23.2-q31.3	TB	68
	17p13.3-p13.1	TB, CA [†]	
	20p13-p12.3	TB, CA	
South Africa	5p15	TST [‡] intensity	63
	11p14	TST positivity	

*TB, current or previous microbiologically confirmed TB

*PTST-, persistently negative tuberculin skin test

[†]CA, ordered subset analysis with minimum age at onset of disease as covariate

[‡]TST, tuberculin skin test

2.3 Association Studies

The most commonly employed study design for the identification of genes that alter susceptibility to TB is candidate gene association studies⁶⁴. This study design is based on the comparison of allele frequencies between cases (affected individuals) and controls (unaffected individuals), provided the alleles are in Hardy-Weinberg equilibrium (HWE)⁶⁹. In essence, association studies involves the investigation of polymorphisms in a gene of interest and whether or not it occurs more frequently in the cases or controls when compared to each other. Association studies are thus hypothesis based⁷⁰. This method of gene identification has greater statistical power compared to linkage analysis, therefore allowing for the identification of genes with smaller effects, provided an adequate sample size is used⁷¹. Another important issue to consider when designing population-based association studies is the type of polymorphisms that will be investigated in the study. It is important to select genes that are involved in the development of TB disease and that the polymorphisms in the candidate gene that are preferably functionally relevant so as to minimize the identification of false-positive associations. Due to the availability of the human genome sequence, selection of functionally relevant polymorphisms has now become possible⁷². Therefore, polymorphisms that result in amino acid changes and thus alter the protein structure are good candidates, in addition to polymorphisms that result in frameshift mutations in the coding area of the gene or alter the expression of the gene⁷². There are usually three reasons why an association is observed between a polymorphism in a candidate gene and the disease of interest, which include (1) the associated allele being the actual cause of the disease, (2) the associated allele is in linkage disequilibrium (LD) with the causative allele of the disease, (3) the association is an artefact of population admixture⁶⁹. It is thus of the utmost importance that significant associations be replicated or validated in other populations and to identify whether or not the population being investigated is stratified, therefore reducing the likelihood of identifying false-positive associations⁷⁰.

There are however disadvantages to this method as well. Firstly, due to the complexity of the disease, it is believed that numerous genes are involved in determining the outcome of the infection. Therefore, using the candidate gene approach could be a very laborious means of identifying these genes. Also, because association studies are hypothesis based, only genes that are known to play a role in immunity against *M. tuberculosis* infection are investigated and this may be problematic as it is possible that many susceptibility genes may not yet have been discovered. This issue can however be addressed by employing genome-wide association studies⁷³. Secondly, when a candidate gene is selected for investigation only a select set of polymorphisms associated with the gene are studied⁷⁴. If these polymorphisms are found not to be associated with the disease it does not necessarily mean that the gene does not play a role in susceptibility to the disease. Thirdly, and probably the biggest concern with regards to association studies, the number of confirmations between studies is low⁷⁰. Associations that are identified in one population are often found not to be associated with the disease in other populations⁷⁵. This is predominantly due to differences in study design between the populations and can include differences in phenotype definition, experimental procedure and sample sizes^{76, 77}. With regards to sample size, more often than not, the initial study will investigate a small subset of the population resulting in reduced statistical power and the identification of false-positives, which is apparent when the initial associations are not found in validation studies using larger sample sizes. Finally, the association of an allele with TB is only a statistical finding and functional experiments are required to identify the biological impact of the associated allele with regards to TB.

There are currently three variations of association studies being employed. These are population-based case-control studies, family-based association studies and genome-wide association studies.

2.3.1 Population-based case-control association studies

Population-based case-control studies are currently the most widely used form of association studies for the identification of genes that alter susceptibility to TB. One of the main advantages of the use of this method, as mentioned previously, is its greater statistical power compared to linkage analysis⁷¹. However, the selection of controls should receive sufficient attention, as controls poorly matched to the cases could negatively affect the results of an association study. Controls and cases should be matched with regards to ethnicity and geographical location. Another confounding factor in association studies is population stratification, as a certain allele or haplotype may be more prevalent in one of the founder populations, which could impact negatively on the power of the study and result in false-positive or false-negative results. To overcome these factors, the use of family based association studies (transmission disequilibrium test, TDT) could be employed⁷⁸. This method investigates the transmission of alleles between heterozygous parents and affected children. Combining these two forms of association studies will yield better results. However, one of the drawbacks of TDT analysis is the requirement of large numbers of families, which is often very difficult. To date, numerous genes and pathways have been studied (Figure 6) to elucidate the host genetic components involved in disease development. Recent association studies identified numerous polymorphisms in these genes as TB susceptibility factors, and some have been successfully validated in other populations (Table 6).

Some of the major genes identified by population-based case-control studies include the human leukocyte antigen (*HLA*) genes, solute carrier family 11A member 1 (*SLC11A1*) (formerly known as natural resistance-associated macrophage protein 1 gene (NRAMP1)) and the pattern recognition receptor mannose-binding lectin (*MBL*) gene.

The *HLA* genes are comprised of approximately 200 genes, and are one of the most extensively investigated gene families⁷⁹. These genes are the most polymorphic in the human genome (3528 alleles) and are predominantly involved in the presentation of antigens to T cells during infection⁸⁰. Several studies investigating the role of polymorphisms in these *HLA* genes in susceptibility to TB have identified numerous alleles which alter TB disease outcome, with most of them highlighting ethnic differences. This is believed to be due to evolutionary selection pressures since the *HLA* genes are involved in the immune response against infectious agents⁷⁹. The *HLA* gene family was one of the first genes to be associated with TB, with the *HLA-DR2* gene having been consistently found to be associated with susceptibility to TB in various populations such as Russia⁸¹, India^{82, 83}, Indonesia⁸⁴ and Thailand⁸⁵.

Solute carrier family 11A member 1 (*SLC11A1*), formerly known as natural resistance-associated macrophage protein 1 (*NRAMP1*) was found to alter susceptibility to leishmania, salmonella and mycobacteria in inbred mouse strains (section 2.3)⁸⁶⁻⁸⁸. Various association studies have been conducted on the role of *SLC11A1* and its associated polymorphisms and susceptibility to TB⁸⁹⁻⁹¹. Numerous genetic variants have thus been identified that alters susceptibility to TB in various populations. A recent meta-analysis of these polymorphisms have shown the 5' (GT)_n variant, D543N (rs17235409) and the 3' UTR (TGTG deletion) variant to be significantly associated with increased risk of pulmonary TB in West African, Asian and South African populations⁹¹.

The mannose-binding lectin (*MBL*) gene encodes the MBL protein, which plays a role in the promotion of phagocytosis and modulation of inflammation^{92, 93}. Various polymorphisms have been found to be associated with susceptibility to TB in various populations. Studies have shown that deficiencies in MBL results in increased susceptibility to various infectious diseases⁹⁴, including TB⁹⁵⁻⁹⁷. This has been hypothesized to be due to the promotion of bacterial uptake into macrophages being advantageous to the bacterium, thus the identification of variant alleles being associated with protection against TB infection⁹⁵.

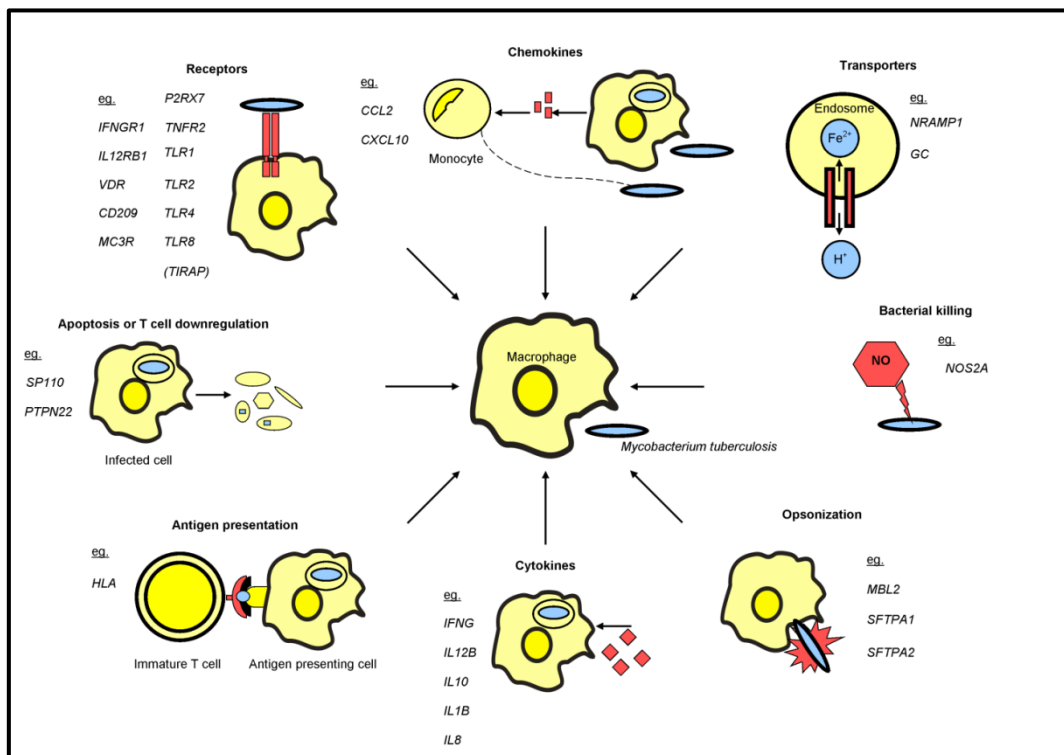


Figure 6: Current understanding of genes involved in altered susceptibility to *M. tuberculosis* infection⁶⁴.

2.3.2 Genome-wide association studies

Due to recent advancements in genotyping technology (microarray chips) and the availability of the human genome sequence and the HapMap Project database⁹⁷⁻¹⁰⁰, genome-wide association studies have become possible. These studies are therefore able to investigate polymorphisms representative of the entire genome, one million single nucleotide polymorphisms (SNPs) in Caucasian populations, irrespective of the genomic location. This method is therefore not subject to the limitations of candidate gene association studies, in which only a few polymorphisms are investigated. With regards to other populations, African for example, a greater number of SNPs is required for complete genome coverage due to greater variation and less LD^{101, 102}. This method is therefore free of assumptions, and allows for the identification of novel genes and pathways.

When conducting GWAS, study design is of the utmost importance. Currently, a two-step design is employed with strict significance cut-off values to limit false-positive associations¹⁰¹. In the first stage of the study, a subset of the study population is investigated on a genome-wide level. Polymorphisms found to be associated with disease in the first stage are then analysed in the rest of the study population. Associated polymorphisms that pass this two-step study design are then validated in a different population in a very large sample size using a different genotyping method to exclude technical artefacts. Associations that remain significant throughout this entire process are then considered as disease susceptibility variants.

There are however draw-backs to this method. Due to the extremely high significant cut-off values employed in this approach to limit the likelihood of false-positive associations, only SNPs that exhibit large effects are identified and all SNPs that demonstrate small to moderate effects are ignored³². GWAS are therefore based on the “common disease, common variant” (CDCV) hypothesis, resulting in all rare variants being missed. There is therefore currently intense debate among researchers with regards to the impact of GWAS due to their basis on the CDCV hypothesis, as some researchers believe that susceptibility to infectious diseases could be better described by a combination of rare variants instead³². GWAS also require extremely large study cohorts to ensure adequate power, because of the number of polymorphisms interrogated¹⁰¹.

GWAS have been successfully employed in identifying susceptible variants in various diseases, including Crohn’s disease, rheumatoid arthritis, diabetes, macular degeneration, inflammatory bowel disease, myocardial infarction and prostate cancer. When it comes to infectious diseases though, the use of GWAS has been limited. There are currently only six published GWAS that have been conducted in infectious diseases; namely HIV¹⁰³, Kawasaki disease¹⁰⁴, chronic hepatitis B¹⁰⁵ and more recently leprosy¹⁰⁶, meningococcal disease¹⁰⁷ and TB⁵⁵. The recently published leprosy GWAS, which was conducted in the Han Chinese population of eastern China, six genes and their associated polymorphisms were found to alter susceptibility to infection by *M. leprae*¹⁰⁸. The genes identified include coiled-coil domain containing 122 (*CCDC122*), chromosome 13 open reading frame 31 (*C13orf31*), nucleotide-binding oligomerization domain containing 2 (*NOD2*), tumour necrosis factor [ligand] superfamily member 15 (*TNFSF15*), *HLA-DR* and receptor interacting serine-threonine kinase 2 (*RIPK2*). The first step of the study was conducted in 706 cases and 1225 controls using the Human610-Quad BeadChip (Illumina), with the 93 associated SNPs identified in the first step replicated in three replication sets (first replication cohort: Han Chinese (eastern China), second and third replication cohort: Han Chinese and non-Han minority ethnic groups) comprising a total of 3254 cases and 5955 controls using the iPLEX (Sequenom) and TaqMan assay (Applied Biosystems) systems. In the field of TB, the only GWAS to be published was conducted in West African populations. This GWAS was a combination of two separate studies

with integrated analysis. The details of this GWAS including the candidate genes investigated will be discussed in chapter 6. A sarcoidosis GWAS, a disease with similar pathomechanisms as TB, was also used to select relevant genes and will be discussed in chapter 5.

2.3.3 Linkage Disequilibrium and Haplotype analysis

Linkage disequilibrium (LD) is the non-random association between two or more alleles at different loci within the genome, whereas haplotypes are a combination of alleles (a DNA sequence) at different chromosomal locations that tend to be inherited together¹⁰¹. In terms of identifying genetic variants that are involved in disease susceptibility, a haplotype would represent a set of SNPs that are statistically associated with the disease. This offers a unique means by which to identify novel susceptibility variants as an associated haplotype block would be indicative of a conserved region of DNA where no recombination occurred, and would contain the causal polymorphism of the disease¹⁰⁹⁻¹¹¹. Currently, haplotypes are predominantly inferred by means of statistical methods due to the high costs associated with laboratory-based haplotyping. However, statistical inferring of haplotypes is not as informative as determining true haplotypes and could result in incorrect haplotype classification^{101, 112}. Importantly, haplotype block structure between populations can differ significantly, as a result of selection, bottle-necks or admixture in the population¹¹¹. Taking this into consideration when comparing haplotype analysis between two populations is essential, as haplotype blocks identified in one population would not necessarily be the same in the other population¹¹³.

The sporadic nature in which mutations occur in the DNA results in the linkage of SNPs along the chromosome⁹⁸⁻¹⁰⁰. Therefore, the presence of one allele at a given SNP could provide information on the alleles present at other polymorphic sites that are in LD and can thus be used for the identification of novel susceptibility SNPs¹¹⁴. LD however also complicates disease susceptibility variant identification, since a SNP which is identified to be associated may in fact not be causal, but is instead in LD with the true disease-causing SNP. This highlights the importance of taking haplotypes into consideration when using LD.

Table 6: Association studies investigating TB susceptibility candidate genes*

Gene	Polymorphism	Population	Phenotype	Cases	Controls	P value (P_c)	OR/RR	Reference
HLA	DQ1	India	pTB	209	122		2.8	84
	DQA	India (South)	pTB	38	36	NS		128
	DQA1*0101	Mexico	pTB	50	95		6.18	129
		Iran	pTB	40	100		2.66	130
	DQA1*0301	Iran	pTB	40	100		0.25	130
	DQA1*0501	Iran	pTB	40	100		0.53	130
	DQA1*0601	Thailand	pTB	82	160	0.02 (0.16)	2.1	85
	DQB	India (South)	pTB	38	36	NS		128
	DQB1*02	Poland	pTB	38	58	0.01	0.39	131
	DQB1*0301	Thailand	pTB	82	160	0.01 (0.13)	0.4	128
	DQB1*0301-0304	South Africa (Venda)	TB	95	117	0.001	2.58	132
	DQB1*0402	Mexico	pTB	65	95		0.18	129
	DQB1*05	Poland	pTB	38	58	(0.002)	2.84	131
	DQB1*0501	Mexico	pTB	50	95		6.16	129
		Iran	pTB	40	100	NS		130
	DQB1*0502	Thailand	pTB	82	160	0.01 (0.13)	2.06	85
	DQB1*0503	Cambodia	pTB	126	88	0.005		133
	DQB1*0601	India (South)	pTB	126	87		2.32	83
	DQB57 Asp/Asp	Cambodia	pTB	436	107	0.001	3.05	134
	DQw1	Indonesia	pTB	101	65		39%	135
	DR2	India	pTB	25 families		0.001		136
		Russia	pTB	135	130	<0.001	3.34	81
		India (South)	pTB	204	404	0.01	0.29	137
		India (North)	pTB	153	289	(0.029)	1.8	138
		India (North)	pTB	124	109	NS		136
		India	pTB	209	122		2.3	84
		Russia (Tuvonian)	pTB	96	291	<0.001 (<0.05)	3.32	139
		Indonesia	pTB	101	65		36%	135
	DR3	Mexico	pTB	51	54	Decreased in cases		140
		Russia	pTB	135	130	<0.05	0.52	81

Gene	Polymorphism	Population	Phenotype	Cases	Controls	P value (P_c)	OR/RR	Reference
	DR4	Mexico	pTB	65	95		0.28	129
		Thailand	pTB	35		Increased in cases		141
		Italy	cTB	54	1089	0.001	2.7	142
	DR8	Mexico	pTB	65	95		0.10	129
	DRB	India (South)	pTB	38	36	NS		128
	DRB1*13	Russia (Tuvinian)	pTB	14		Over transmitted		143
				pedigrees				
	DRB1*07	Iran	pTB	40	100		2.7	130
	DRB1*0803	Korea	DR pTB	81	200	0.047	2.63	144
	DRB1*11	China	pTB	74	90	<0.05	0.12	145
	DRB1*13	Poland	pTB	31	58	<0.001	0.04	146
	DRB1*1302	South Africa (Venda)	TB	92	117	<0.001	5.05	132
	DRB1*14	Russia (Tuvinian)	pTB	14		Over transmitted		143
				pedigrees				
	DRB1*15	China	pTB	74	90	<0.05	2.91	145
	DRB1*1501	Mexico	pTB	50	95		7.92	129
		India	pTB	22	36	<0.05		147
		India (South)	pTB	126	87		2.68	83
	DRB1*16	Poland	pTB	31	58	<0.01	9.7	146
	DRw53	Russia (Tuvinian)	pTB	96	291	<0.001 (<0.05)	11.88	139
	HLA	Brazil	pTB	98		NS		148
				pedigrees				
	HLA-A, -B, -C	Italy	TB	68	1089	NS		142
		India (North)	pTB	124	109	NS		136
	HLA-A and -B	China (Hong Kong)	pTB	256	100	NS		149
CCL2	-2518	China (Hong Kong)	TB	412	456	NS		150
		South Africa (SAC)	pTB	431	482	NS		151
		Russia	pTB	1440	1529	NS		152
		Ghana	pTB	2010	2346	(1.8×10^{-3})	0.81	152
		Mexico	pTB	445	334	0.0003	2.43	153

Gene	Polymorphism	Population	Phenotype	Cases	Controls	P value (P_c)	OR/RR	Reference
		Korea	pTB	129	162	0.0001	2.63	153
	-362C	Ghana	pTB	2010	2346	(2.3×10^{-4})	0.83	152
CD209	-336	Columbia	TB	110	229	NS		154
		Tunisia	TB	138	140	NS		155
		Guinea-Bissau	pTB	321	347	NS		156
		India (South)	TB	107	157	NS		157
		South Africa (SAC)	pTB	351	360	0.01	1.48	158
		Gambia	TB	329/347	327	0.01/0.03	0.75/0.79	159
		Republic of Guinea	TB	151	180	NS		159
		Guinea-Bissau	TB	162	141	NS		159
		Malawi	TB	244	295	NS		159
		sub-Saharan Africa	TB	1233	943	0.006	0.86	159
		South Africa (SAC)	pTB	351	360	8.2×10^{-4}	1.85	158
CXCL10	-1447	China	TB	240	176	NS		160
		China	TB	240	176	NS		160
		China	TB	240	176	0.01	0.51	160
GC	Gc	Brazil	TB	130	78	NS		161
		South Africa	TB	281	182	NS		161
		UK (Gujarati Asian)	TB	123	140	0.009	2.81	161
IFNG	+874	Sicily	pTB	45	97	0.02		162
		Croatia	TB	54	175	0.012	3.12	163
		Caucasian	pTB	113	207	0.0017	3.75	164
		West Africa	TB	682	619	NS		165
		China	TB	301	310	0.035	1.98	166
		South Africa (SAC)	pTB	313	235	0.0055	1.64	167
		Turkey	pTB	319	115	0.024	0.7	168

Gene	Polymorphism	Population	Phenotype	Cases	Controls	P value (P_c)	OR/RR	Reference
		India (South)	TB	129	127	NS		169
		African American	TB	240	174	NS		170
		Caucasian	TB	161	64	NS		170
		USA (Houston)	TB	319	98	NS		170
		China (Hong Kong)	pTB	385	451	<0.001	2.24	171
		Croatia	pTB	253	519	NS		163
		Columbia	TB	190	135	0.01		172
		Meta-analysis	TB	11 studies		0.0008	0.75	173
	1348T/A	Japan	pTB	114	110	0.55		174
	CA repeat	Indonesia	TB	382	437	NS		175
	G2109A	Croatia	pTB	253	519	NS		163
	rs2069718	Japan	TB	87	265	NS		176
	rs2193049	Japan	TB	87	265	NS		176
<i>IL10</i>	-1082	Ghana	pTB	2010	2346	NS		177
		Pakistan	pTB	111	188	NS		178
	-2849A/-1082A/-819C/-592C	Ghana	TST	2010	129 TST- 2219TST+	0.013 0.017	2.15 2.09	177
<i>IRF1</i>	5 SNPs	Vietnam	pTB	162	132	NS		179
	17 SNPs	Indonesia	pTB	192	192	NS		179
<i>MBL</i>	A/O	Tanzania	TB	443	426	NS		180
	A/O and O/O	Spain	TB	106	344	0.02	0.58	181
	B	South Africa (SAC)	TBM	91	79	0.017		95
		African-American	pTB	176	71	<0.01	0.34	182
	C	Gambia	pTB	397	422	0.037	0.79	183
	C, D	African-American	pTB	176	71	NS		182
	codon 54	Turkey	TB	44	99	0.01	3.9	96
	codons 52, 54 and 57	Poland	pTB	126	124	NS		184
	HYA/HYA	Italy	pTB	277	288	1 x 10⁻⁸	0.09	185

Gene	Polymorphism	Population	Phenotype	Cases	Controls	P value (P_c)	OR/RR	Reference
	LYB/LYD	Italy	pTB	277	288	1 x 10⁻⁶	49	185
	O	India	pTB	202	109	0.008	6.5	186
		USA (Houston)	pTB	198	46	NS		182
		Caucasian	pTB	113	69	NS		182
		Spain	TB	106	344	0.009	0.6	181
	O, H, L, X, Y, P, Q	China	pTB	152	293	NS		187
	O, X	Denmark	pTB	59	250	0.03		97
	X/O and O/O	Spain	TB	106	344	0.023	0.46	181
<i>NOS2A</i>	CCTTT	Colombia (Paisa)	pTB	114	304	0.0001 (0.001)	0.4	188
	rs1800482	Mexico	pTB	445	518	NS		153
	rs2274894	USA (African-American)	pTB	279	166	0.003	1.84	189
	rs2779249	Brazil	pTB	92		0.021	3.25	189
	rs7215373	USA (African-American)	pTB	279	166	0.004	1.67	189
	rs9282799 and rs8078340	South Africa (SAC)	pTB	431	482	0.011 (0.029)	1.4	151
	TAAA	Colombia (Paisa)	pTB	114	304	NS		188
								190
<i>P2RX7</i>	-1513	China (Han Chinese)	TB	96	384	NS		191
		Gambia	TB	>300	>160	NS		192
		Liverpool	pTB	86	167	NS		193
		Mexico	TB	94	100	0.02	5.28	194
		Russia	pTB	190	128	0.02	1.71	192
		Sydney	pTB	99	102	NS		190
	-762	China (Han Chinese)	TB	96	384	NS		191
		Gambia	TB	>300	>160	0.003	0.70	193
		Mexico	TB	94	100	NS		194
		Russia	pTB	190	128	NS		190

Gene	Polymorphism	Population	Phenotype	Cases	Controls	P value (P_c)	OR/RR	Reference
PTPN22	R263Q	Morocco	pTB	123	155	0.01	5.85	195
	R620W	Morocco	pTB	123	155	0.01	0.14	195
SLC11A1	1703G/A	Japan	pTB	114	110	0.144		174
	274C/T	USA (Houston)	Pediatric TB	184		0.01	1.75	196
	3' UTR	Thailand	TB	149	147	NS		197
		Morocco	pTB	116		NS		198
				pedigrees				
		China	TB	278	282	0.0165	1.51	199
		Korea	pTB	192	192	0.02	1.85	200
		China	Pediatric TB	136	435	Associated		201
	4 SNPs	Japan	TB	87	265	NS		176
	5' CA	Tanzania	TB	443	426	0.014	1.45	180
	ATA _n	Gambia	pTB	318	146	NS		202
	D543N	Thailand	TB	149	147	NS		197
		China	TB	278	282	NS		199
		China	pTB	61	122	NS		203
		Cambodia	pTB	358	106	0.02	0.59	204
		Gambia	pTB	410	417	0.004	1.85	89
		China (Han Chinese)	pTB	110	180		1.93	205
		China	Severe TB	127	91		2.27	206
		China (Han Chinese)	pTB	120	240		2.59	207
		Japan	cTB	95	90		5.16	208
		Peru	pTB	507	513	<0.05	1.4	209
	GT _n	South Africa (SAC)	pTB	265	224	0.002	0.6	90
		Morocco	pTB	116		NS		198
				pedigrees				
		Poland	pTB	85	93	0.03	1.69	198
		USA (Houston)	Pediatric TB	184		0.04		196

Gene	Polymorphism	Population	Phenotype	Cases	Controls	P value (P_c)	OR/RR	Reference
				families				
		Gambia	pTB	329	324	0.024	1.4	210
		Gambia	pTB	410	417	0.008	1.45	89
		USA (Caucasian)	pTB	135	108		2.02	211
		Japan	pTB	202	267	0.0003	2.07	212
	INT4	Russia (Slavic)	pTB	58	127	NS		213
		Taiwan	pTB	49	48	NS		214
		Guinea-Conakry	pTB	44 families		0.036		215
		Poland	pTB	126	124	NS		184
		Gambia	pTB	410	417	0.006	1.83	89
		Denmark	TB	104	176	0.013	1.9	97
		China	Severe TB	127	91		2.29	206
		China	pTB	61	122		2.73	203
		China	Pediatric TB	136	435	NS		201
		Peru	pTB	507	513	<0.05	1.72	209
	TGTG+/d	Gambia	pTB	410	417	0.004	1.85	89
		China (Han Chinese)	pTB	147	145	<0.01		216
		Cambodia	pTB	358	106	0.02	0.59	204
		China (Han Chinese)	pTB	120	240		0.89	207
		China (Han Chinese)	pTB	110	180		2.22	205
		South Africa (SAC)	pTB	265	224	0.013	5.19	90
<i>SP110</i>	rs2114592	Republic of Guinea	pTB	99 families		0.009		217
		Guinea-Bissau	pTB	102		0.002		217
		Gambia	pTB	219 families		0.02		217
		South Africa (SAC)	TB	381	417	NS		218
	rs3948464	Republic of Guinea	pTB	99 families		0.015		219
		Guinea-Bissau	pTB	102		NS		220
				families				

Gene	Polymorphism	Population	Phenotype	Cases	Controls	P value (P_c)	OR/RR	Reference
		Gambia	pTB	219 families		0.01		217
		South Africa (SAC)	TB	381	417	NS		217
		Russia	pTB	1912	2104	NS		219
		Ghana	pTB	2004	2366	NS		217
	rs41547617	Republic of Guinea	pTB	99 families		NS		217
		Guinea-Bissau	pTB	102 families		NS		217
		Gambia	pTB	219 families		0.02		217
<i>TIRAP</i>	rs8177374	Gambia, Guinea-Bissau, Republic of Guinea	TB	675	605	0.04 (0.013)	0.23	221
		Columbia	TB	147	391	0.03	0.53	222
		Russia	pTB	1867	2076	NS		223
		Indonesia	pTB	611	681	NS		223
		Ghana	pTB	1913	2293	NS		223
		Vietnam	pTB	183	392	NS		224
		Vietnam	TBM	175	392	0.001	3.02	224
<i>TLR1</i>	248S	USA (African-American)	TB	339	194	0.009	1.63	225
		USA	TB	99 families		0.021		225
	S602I	USA (African-American)	TB	339	194	<0.001	2.50	225
		USA	TB	99 families		0.021		225
<i>TLR2</i>	(GT) _n < 16	Korea	TB	176/164	196	0.047/0.02	1.41	226
	Arg753Gln	Guinea-Bissau	pTB	321	347	NS		156
		Turkey	pTB	151	116	0.022	6.04	227
	Ins/Del (-196 to -174)	Guinea-Bissau	pTB	321	346	0.023	0.70	228

Gene	Polymorphism	Population	Phenotype	Cases	Controls	P value (P_c)	OR/RR	Reference
		USA (African-American)	pTB	295	179	NS		228
		USA (Caucasian)	pTB	237	144	0.0007	0.41	228
<i>TLR4</i>	Asp299Gly	Gambia	pTB	307	298	NS		229
		Guinea-Bissau	pTB	321	347	NS		156
<i>TLR8</i>	rs3764880	Indonesia	pTB	375	387	0.007	1.8	61
		Russia	pTB	1837	1779	0.03	1.2	
<i>TLR9</i>	rs352143	Guinea-Bissau	pTB	321	346	NS		228
		USA (African-American)	pTB	295	179	0.029	0.58	228
		USA (Caucasian)	pTB	237	144	0.017	0.53	228
	rs5743836	Guinea-Bissau	pTB	321	346	NS		228
		USA (African-American)	pTB	295	179	0.014	0.63	228
		USA (Caucasian)	pTB	237	144	0.05	0.58	228
<i>TNFRSF1B</i>	rs3397	South Africa (SAC)	TB	429	482	0.049	1.22	230
		Ghana	TB	640	1158	0.007	1.315	230
<i>VDR</i>	Various SNPs	Japan	TB	87	265	NS		176
	Apal	Tanzania	TB	443	426	NS		180
		South Africa (Venda)	TB	~85	~88	NS		132
		Guinea-Bissau	pTB	321	347	0.03		156
	BsmI	South Africa (Venda)	TB	~85	~88	NS		132
		India	Spinal TB	64	103		2.2	231
		Asia (Gujarati)	eTB	52	116	NS		232
	FokI	Tanzania	TB	443	426	NS		180
		South Africa (Venda)	TB	~85	~88	NS		132

Gene	Polymorphism	Population	Phenotype	Cases	Controls	<i>P</i> value (<i>P_c</i>)	OR/RR	Reference
		India	Spinal TB	64	103		2.4	231
		Asia (Gujarati)	eTB	52	116		2.8	232
		China (Han Chinese)	pTB	76	171		3.67	233
		China (Han Chinese)	pTB	120	240	0.03	2.35	207
		Peru	TB treatment	103	206		9.6	234
	FokI-BsmI-ApaI-TaqI	West Africa	pTB	382 trios		0.009		235
	TaqI	Tanzania	TB	443	426	NS		180
		South Africa (Venda)	TB	~85	~88	NS		132
		Cambodia	pTB	358	106	NS		204
		Asia (Gujarati)	eTB	52	116	NS		232
		Peru	TB treatment	103	206		5.6	234
		Gambia	pTB	408	414	0.01	0.53	236

*Adapted from ³² and ⁶⁴. Significant associations are in bold.

P_c, corrected *P*-value; OR, odds ratio; RR, relative risk; pTB, pulmonary TB; cTB, cavitory TB; eTB, extrapulmonary TB; TBM, TB meningitis; TST, tuberculin skin test; NS, not significant.

2.4 Animal Models

There are currently numerous animal models; such as the fish, guinea-pig, mouse and rabbit TB models, which are being used to understand disease susceptibility¹¹⁵. These models provide several advantages, in that one can control their breeding as well as the influence of environmental factors. Mouse models are however the most frequently used animal model, especially due to advances in molecular biology, allowing the creation of 'knock-out' (KO) mice, which lack genes that are important in the immune system. KO mice can illustrate the importance or redundancy of genes thought to alter susceptibility to TB disease.

In the 1970's, the natural resistance associated macrophage protein (*Nramp1*), was identified in a mouse model of mycobacterial disease by positional cloning, as a gene that resulted in resistance to infection by various pathogens (see *section 2.3.1*)¹¹⁶. Today, the human homologue, *NRAMP1*, is considered to be one of the first genes found to alter susceptibility to TB in humans¹¹⁷. By using the KO method, various genes involved in the immune response to *M. tuberculosis* infection in the mouse model have been targeted, and numerous genes shown to result in increased susceptibility to various mycobacterial species have been identified (Table 6).

Table 7: Murine genes which result in increased susceptibility to mycobacteria species when disrupted

Gene	Increased susceptibility to	Reference
Nramp1	Salmonella, Leishmania and mycobacteria species	116
55 kD TNF receptor	<i>M. bovis</i> (BCG)	118
IFN- γ	<i>M. tuberculosis</i>	119, 120
IFN- γ Receptor	<i>M. bovis</i> (BCG)	121
β 2-microglobulin	<i>M. tuberculosis</i> / <i>M. bovis</i> (BCG)	122, 123
MHC class II	<i>M. bovis</i> (BCG)	123
T cell receptor	<i>M. bovis</i> (BCG)	124
INF regulatory factor 1	<i>M. bovis</i> (BCG)	125
IL-6	<i>M. tuberculosis</i>	126
IL-12	<i>M. tuberculosis</i>	127

Chapter 3: Hypothesis and Aims

That is the essence of science: ask an impertinent question, and you are on the way to the pertinent answer.

Jacob Bronowski

3.1 Study Hypothesis

Background: A genome-wide association study is a useful technique for the identification of novel genes and pathways that could alter susceptibility to disease. Recent GWAS investigating various granulomatous diseases have identified such novel variants which could also hold true for susceptibility to TB.

1. A GWAS conducted in a German population to identify novel susceptibility variants for sarcoidosis identified *ANXA11* polymorphisms (Hofmann et al, 2008). Sarcoidosis shares various clinical features with TB pathogenesis, which makes susceptibility variants identified, suitable candidates for TB.
2. The first GWAS in TB was conducted in a West African population and identified a novel susceptibility gene, *CADM1* (unpublished data). A validation of these results was attempted in our South African Coloured (SAC) population to determine if the genes were also associated with susceptibility to TB in our population.

Hypothesis: Genes identified in GWAS for sarcoidosis (ANXA11) and TB (CADM1) in other populations will be associated with TB in the SAC population.

3.2 Study Aims

1. To investigate whether polymorphisms found to be associated with granulomatous diseases in other populations are associated with susceptibility to TB in the SAC population, using a case-control association study.
2. To determine LD patterns and identify haplotypes in the candidate genes.
3. To functionally characterize any significant association identified in the case-control association studies.
4. To understand if the methylation pattern of *ANXA11* could influence TB disease outcome.

Chapter 4: Study Methods

The most exciting phrase to hear in science, the one that heralds new discoveries, is not Eureka! (I found it!) but rather, "hmm.... that's funny...."

Isaac Asimov

4.1 Reagents and Equipment

All reagents and equipment used in this study are listed in the addendums 1, 2 and 3.

4.2 Study Participants

4.2.1 Study Population

The participants of this study belong to the South African Coloured (SAC) population, which is a highly admixed population with genetic contributions from African Blacks, Europeans, Khoi, San and Indonesia^{79, 237}. Due to this “mixing” of different ethnicities to form the SAC population, the presence of population stratification was possible, which could affect the results (false positives/negatives) obtained during association studies due to differential allelic-expression within each of the founder populations²³⁸.

To test whether the SAC population was stratified, a study in which 25 unlinked SNP markers were genotyped was undertaken in this population¹⁵⁸. No significant stratification ($P = 0.26$) was detected. Using admixed populations in association studies is highly advantageous due to the presence of greater number of alleles within members of this population, since their genetic make-up is a combination of the ancestral populations. This therefore affords us the opportunity to investigate unique markers with disease association within a single population²³⁹.

4.2.2 Case-Control Samples

The individuals recruited for this study were from the Ravensmead and Uitsig suburbs of the Western Cape province of South Africa, with approximately 98% of residents self-identifying themselves as SAC²³⁷. The Western Cape has one of the highest incidences of TB in the country, with 967 per 100 000 in 2006²⁴⁰. Individuals were therefore selected from these suburbs due to the high incidence of TB but low prevalence of HIV (2%)^{241, 242}.

All study participants included in this study were unrelated. The case samples consisted of individuals who had bacteriologically (smear positive and/or culture positive) confirmed pulmonary TB. The control samples were healthy individuals with no history of TB but lived in the same community, and were at least 17 years of age. All study participants were HIV negative. For the characteristics of the case-control samples, see table 8.

Table 8: Characteristics of individuals included in the case-control association studies.

Characteristic	Cases	Controls
Pulmonary TB Cohort		
Males: n (%)	200 (48%)	107 (26%)
Females: n (%)	217 (52%)	303 (74%)
Average age (years \pm SD)	33 \pm 13.6	32 \pm 10.8
Total number	417	410
TB Meningitis Cohort		
Males: n (%)	46 (43%)	17 (23%)
Females n (%)	61 (57%)	57 (77%)
Average age (years \pm SD)	7 \pm 5.2	29.7 \pm 7.9
Total number	107	74

n, number; SD, standard deviation

4.3 DNA Samples

4.3.1 DNA Extractions

Blood samples were collected from individuals with informed written consent and approved by the Health Research Ethics Committee of Stellenbosch University, South Africa (Project registration number 95/072). DNA was extracted and purified using the Nucleon BACC Genomic DNA Extraction Kits (Illustra, Buckinghamshire, UK) following the manufacturer's instructions.

Once the DNA was extracted, concentration and purity was determined using the NanoDrop® ND-1000 Spectrophotometer and the NanoDrop® v3.0.1 Software (Inqaba Biotechnology, Pretoria, SA). Working solutions were prepared for general PCR at a concentration of 100 ng/μl and stored at -20°C, while the DNA used for the TaqMan® genotyping system (Applied Biosystems, Foster City, USA) was diluted to a concentration of 20 ng/μl and stored in 96 deep-well plates (Eppendorf, Hamburg, Germany) at -20°C. DNA was diluted with milli-PORE (Massachusetts, USA) water.

4.3.2 Plate Design

The 96 well plate layouts were imported into an Excel file, with each plate containing 84 DNA samples and 12 negative controls (distilled water). These plates were used for TaqMan® genotyping. For high-throughput genotyping 384 well plates were employed. Four 96 well plates were merged into one 384 well plate, with sample allocation determined using the Excel 7900_96 to 384 well template (Figure 7). Each 384 well plate received a unique name which identified the polymorphism being investigated and the samples contained within it, example rs7071579A – rs7071579C.

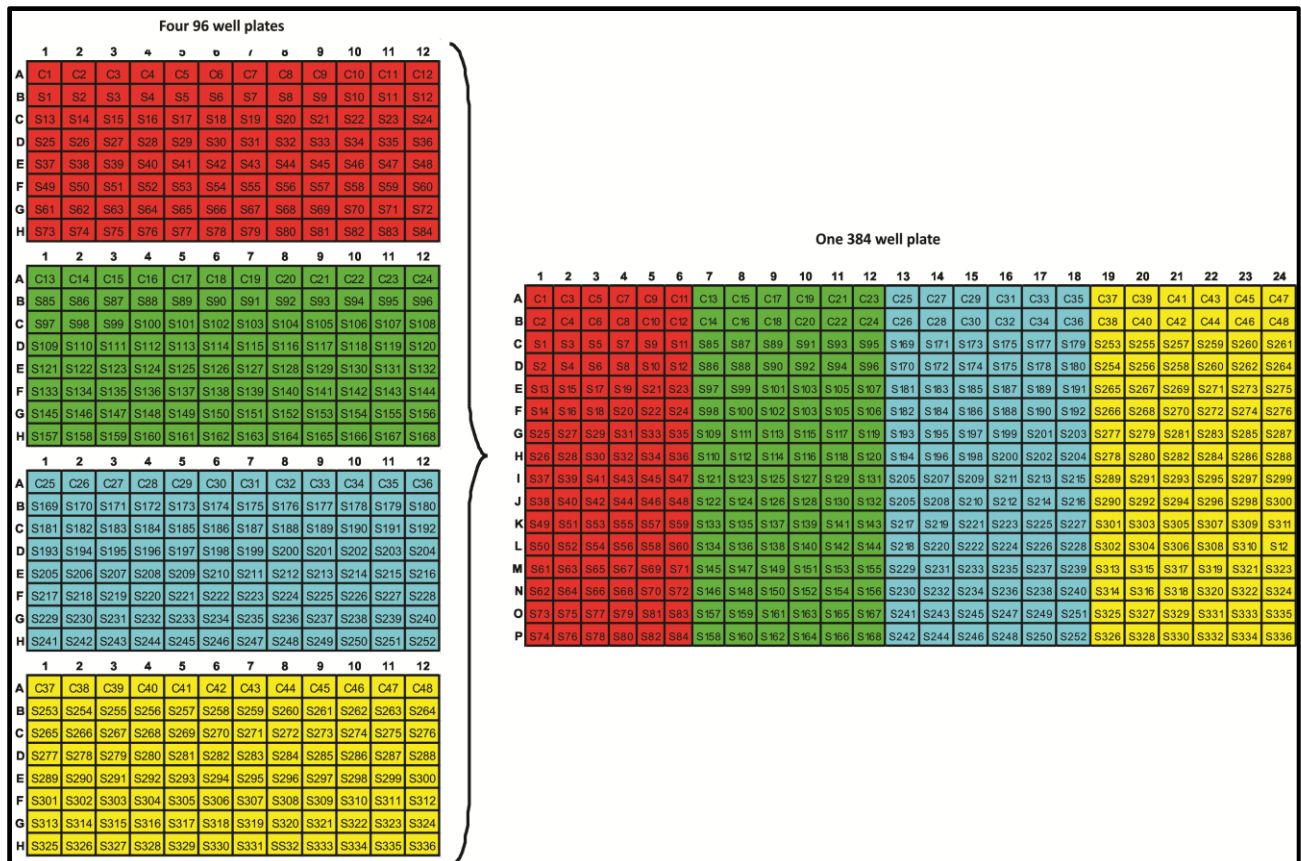


Figure 7: Schematic representation of sample allocation in 384 well plates.

4.4 Genotyping Methods

4.4.1 TaqMan® Genotyping System

TaqMan® SNP genotyping is a highly efficient method for determining allelic variation. It is a one-step PCR reaction (Figure 8) which relies on the 5' to 3' exonuclease activity of *Taq* DNA polymerase²⁴³. The basic principle behind the TaqMan® genotyping system is quite simple. Firstly, a primer set is designed that is specific for the polymorphism and two probes which are specific for each allele (Figure 9). These probes contain a fluorescent dye at the 5' end and a quencher molecule at the 3' end, and as long as these two molecules are in close proximity with each other the fluorescence emitted by the dye is diminished²⁴⁴. The 3' end also has a minor groove binder attached which is responsible for the stability of the double-stranded probe structure. This results in an increase in the melting temperature (T_m) of the probe without increasing its length, augmenting the allelic discrimination capabilities of the probe. Each probe also contains a different fluorescent dye, which allows for the discrimination of the two alleles in a single reaction²⁴⁵.

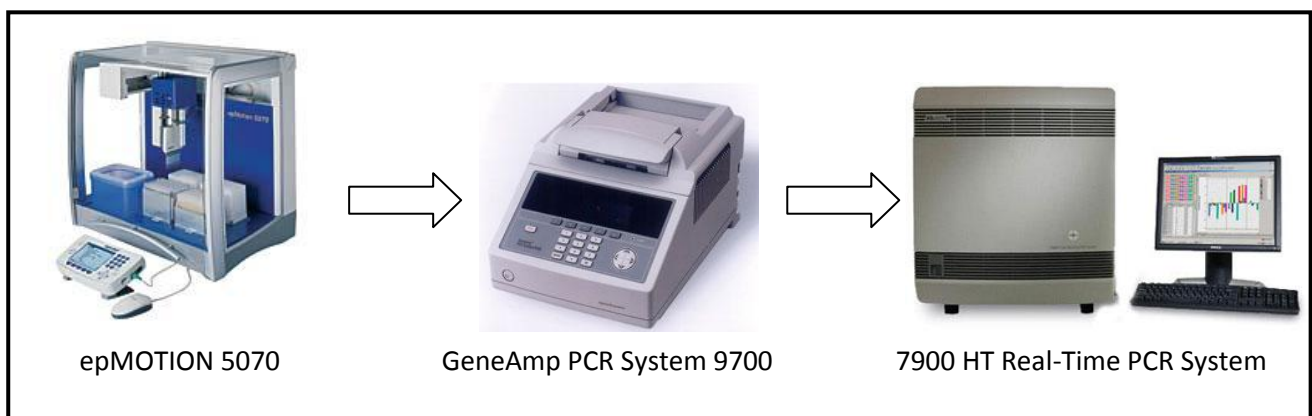


Figure 8: Schematic representation of the TaqMan® Genotyping System Flowthrough.

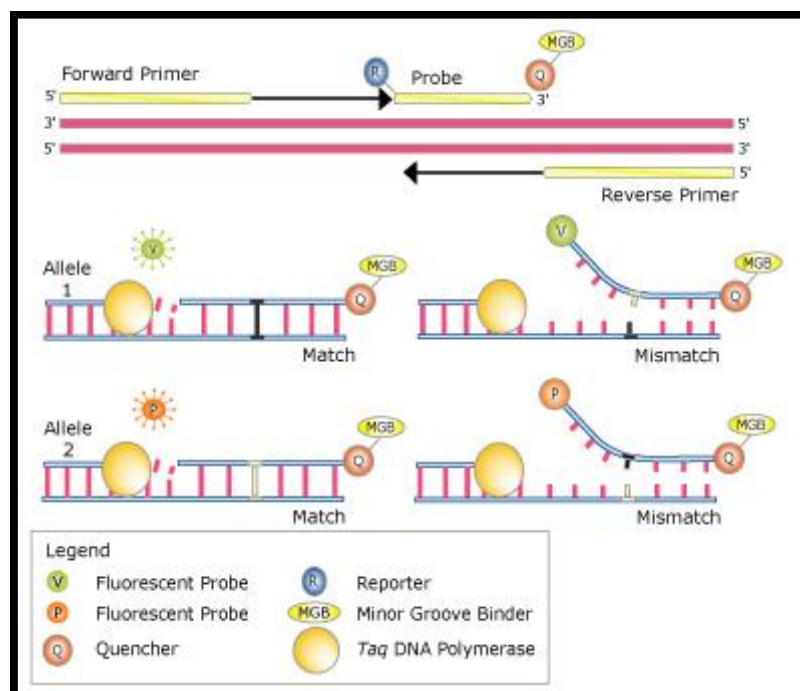


Figure 9: Graphical representation of the TaqMan® Genotyping Assay²⁴⁶.

During the PCR reaction, the primers and probes are able to anneal to the template DNA (Figure 9). If the probe recognizes the allele it will bind to the DNA allowing for the 5' exonuclease activity of the *Taq* polymerase to cleave the fluorescent dye from the probe, allowing it to give off a fluorescent signal. However, if the probe is not specific for the allele it will not bind resulting in the fluorescent dye remaining in close proximity to the quencher molecule. Once the PCR reaction is complete, genotype calling is done by plotting the normalised intensity of the fluorescent dye on a cluster plot (Figure 10) using data analysis software. For diallelic polymorphisms, three such clusters are possible, with two clusters representing the homozygote genotypes (emission of only one of the dyes) and the third cluster representing the heterozygote genotype (emission of both dyes in the same sample). Finally, a fourth cluster is also identified and representative of the negative control samples, and usually occurs closest to the x- and y- axis intersection.

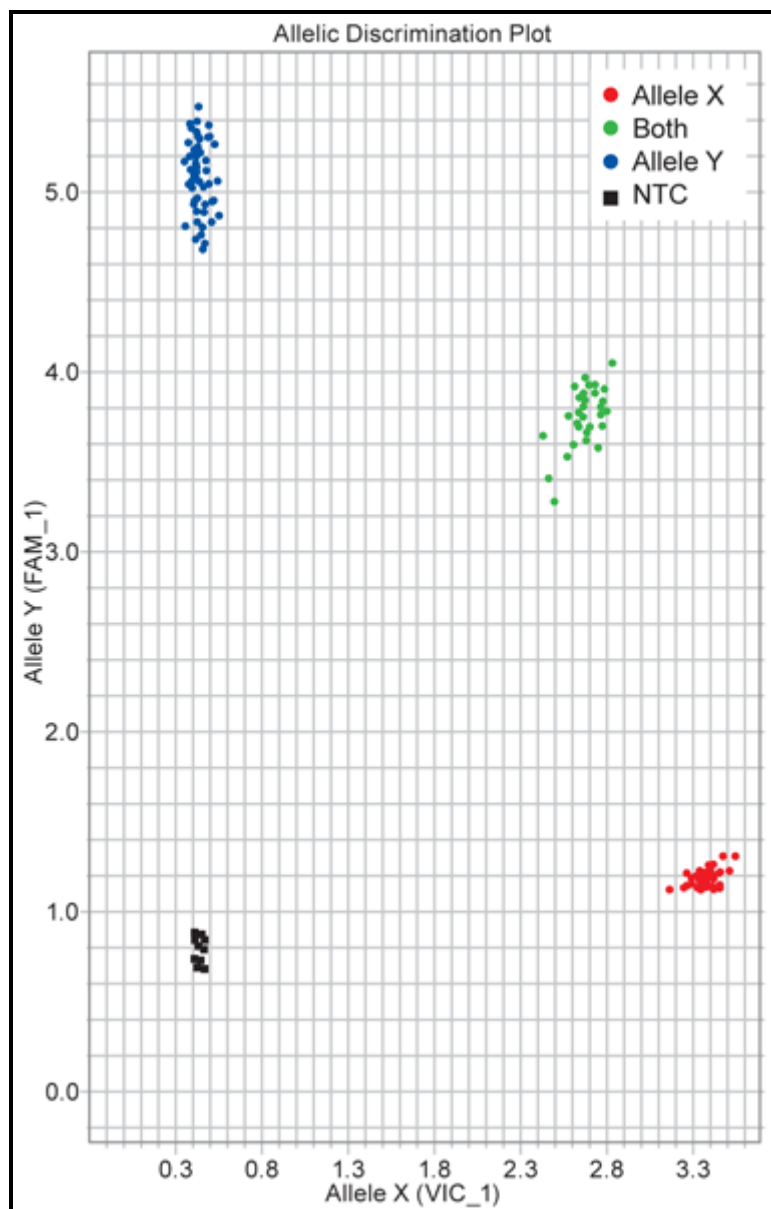


Figure 10: Cluster plot used for the calling of genotypes²⁴⁷.

4.4.2 TaqMan® Protocol

In this study, five polymorphisms of *CADM1*, one polymorphism of *CADM2*, three polymorphisms of *CADM3*, one polymorphism of *NCAM2* and six polymorphisms of *ANXA11* were genotyped using the TaqMan® Genotyping system (Table 9). Pipetting was done using the Eppendorf epMotion 5070 automated pipetting system. For each 5 µl reaction, 2.5 µl TaqMan® Genotyping Master Mix (Applied Biosystems), 0.06 µl (40X) of TaqMan® SNP Genotyping Assay (Applied Biosystems) and 2.4 µl of distilled water was added. To this master mix, 20 ng of DNA was added. PCR reactions were completed in the GeneAmp® PCR System 9700 (Applied Biosystems) machine with the following cycling conditions: 10 minutes at 95°C, followed by 45 cycles of 10 seconds at 95°C and 1 minute at 60°C. The plates were analysed on the 7900 HT Fast Real-Time PCR System (Applied Biosystems) using the SDS v2.3 software (Applied Biosystems).

Table 9: TaqMan® Genotyping Assays used in this study.

Gene	SNP	Assay ID
<i>CADM1</i>	rs1460911	C_1683071_10
	rs1563899	C_7487749_10
	rs1563900	C_7487750_10
	rs2446890	C_1683096_10
	rs2515327	C_26488323_10
<i>CADM2</i>	rs2324979	C_7724485_10
<i>CADM3</i>	rs12057331	C_31622088_10
	rs12068892	C_31622098_10
	rs16841729	C_33822427_10
<i>NCAM2</i>	rs8184921	C_29840603_10
<i>ANXA11</i>	rs2789679	C_16108006_10
	rs7071579	C_1817519_10
	rs1953600	C_12123137_10
	rs2573346	C_1817475_10
	rs2784773	C_1817509_10
	rs1049550	C_7881261_10

4.5 SNP Identification (*ANXA11*)

We were also interested in identifying novel polymorphisms or other potential disease associated polymorphisms in *ANXA11*. We therefore sequenced 20 individuals, 10 cases and 10 controls, at various regions within and flanking the gene. We were interested in identifying polymorphisms that could be in LD with the rs1049550 (exon 5, causative allele in sarcoidosis GWAS) and the rs7071579 polymorphism (3' UTR) which was associated with susceptibility to TB in the SAC population. Polymorphisms that occurred within a functional region (promoter, exon, exon-intron splice site) of the gene and had a minor allele frequency of at least 5% were subsequently investigated as a possible TB susceptibility variant in our SAC population.

All sequencing reactions were conducted by the Central Analytical Facility (Stellenbosch University, South Africa), the results of which were aligned using the Sequencher v4.7 software (Gene Codes Corporation, Michigan, USA). All variations identified were analysed further using the web based program Blat (<http://genome.ucsc.edu/cgi-bin/hgBlat>). After inserting the flanking sequence of the polymorphism, Blat identified all the known polymorphisms within that region with their respective rs-numbers. The rs-number represents a unique ID that can be used to search the Single Nucleotide Polymorphism Database (dbSNP) which is hosted by the National Human

Genome Research Institute and the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/projects/SNP/>). If no rs-number was provided and there was no reference to the SNP in the literature, it was considered to be novel.

Four regions spanning *ANXA11* were selected for sequencing, including exons 4, 5 and 6 and the 3' UTR (Table 10).

Table 10: Primer sets and their respective sizes used to sequence *ANXA11* regions.

Name	Sequence (5' – 3')	Size (bp)
E4seqF	TCAAATCTGGGATATGGCTCTT	222
E4seqR	ACGTACACACCTTGATGGCTTC	
E5-6seqF	CGGCATGCACGACATCTTACCT	468
E5-6seqR	GCCTTGCCGTAAGCCGTCTTGA	
3UTRseqF	TGCTGAAGATCTGTGGTGGCAATGAC	747
3UTRseqR	ACCATTCCAGAGCTGGCTGACTCTT	

Primers were manufactured by Integrated DNA Technologies (Iowa, USA).

4.5.1 Exons 4, 5 and 6

Exons 5 and 6 were sequenced in one reaction while exon 4 was sequenced on its own. Each sample was amplified in a total volume of 25 µl containing: 2.5 µl 10X Buffer (containing 1.5 mM MgCl₂) (Southern Cross Biotechnologies, Cape Town, South Africa), 2 µl 1.25 mM dNTPs (Bioline, London, UK), 10 µM of each primer, 17.425 µl distilled water, 0.075 µl SuperTherm GOLD HotStart *Taq* polymerase and 100ng of DNA. Cycling conditions were as follows: 10 minutes at 94°C, followed by 35 cycles of 30 seconds at 94°C, 30 seconds at 60.6°C (exon 7 and 8)/ 54.6°C (exon 9) and 45 seconds at 72°C, followed by an extension step of 10 minutes at 72°C and a cooling step of 2 minutes at 4°C.

4.5.2 3' UTR

The 3' UTR region was amplified in a total volume of 25 µl containing: 2.5 µl 10X Buffer (containing 1.5 mM MgCl₂) (Southern Cross Biotechnologies), 2 µl 1.25 mM dNTPs (Bioline), 10 µM of each primer, 17.425 µl distilled water, 0.075 µl SuperTherm GOLD HotStart *Taq* polymerase and 100ng of DNA. Cycling conditions were as follows: 10 minutes at 94°C, followed by 35 cycles of 30 seconds at 94°C, 30 seconds at 61.2°C and 45 seconds at 72°C, followed by an extension step of 10 minutes at 72°C and a cooling step of 2 minutes at 4°C.

4.5.3 PCR Clean-Up (*ExoSAP-IT*)

All amplified samples were electrophoresed on a 1.5% agarose gel at 150 V for 30 minutes; image capturing was done using the G-Box (Syngene, Cambridge, UK). All samples that were successfully amplified and contained no spurious amplification were subjected to a PCR clean-up using *ExoSAP-IT* (USB Corporation, Ohio, USA). The clean-up reaction was done in a 16.8 µl reaction volume, containing 4.8 µl *ExoSAP-IT* and 12 µl PCR product. Cycling conditions for the clean-up process were as follows: 15 minutes at 37°C, 15 minutes at 80°C and 2 minutes at 4°C. Primers for sequencing were 1.1 µM, with each sequencing reaction requiring 5 µl. These primers and cleaned-up PCR products were sent for sequencing at the Central Analytical Facility (Stellenbosch University). Sequencing results were analysed using Sequencher v4.7 (Gene Codes Corporation, Michigan, USA).

4.6 Gene Expression Analysis

4.6.1 Sample Selection and Genotyping

Heparinized venous blood was obtained from healthy volunteers after informed consent and approval by the Health Research Ethics Committee of Stellenbosch University, South Africa (Project registration number 10/08/249). These DNA samples were sequenced to determine the genotype for each individual at the rs7071579 polymorphic site. Individuals who were homozygous for the A- or G-allele were then selected to be part of the study.

A primer set (Table 11) was designed to genotype the DNA samples. Each sample was amplified in a total volume of 25 μ l containing: 2.5 μ l 10X Buffer (containing 1.5 mM MgCl₂) (Southern Cross Biotechnologies), 2 μ l 1.2 5mM dNTPs (Bioline), 10 μ M of each primer, 17.425 μ l distilled water, 0.075 μ l SuperTherm GOLD HotStart *Taq* polymerase and a 100ng of DNA. Cycling conditions were as follows: 10 minutes at 94°C, followed by 35 cycles of 30 seconds at 94°C, 30 seconds at 52.3°C and 45 seconds at 72°C, followed by an extension step of 10 minutes at 72°C and a cooling step of 2 minutes at 4°C.

For sample visualization and clean-up, see *section 4.5.3*.

Table 11: Primer set used to sequence the rs7071579 polymorphism.

Name	Sequence (5' – 3')	Size (bp)
rs7071579F	GAGATCTAGAAAGGCCAGTCCATGCTACACATC	286
rs7071579R	GAGAGGATCCGACTCGGTCTTGTCGGTTTCAGG	

Primers were manufactured by Integrated DNA Technologies.

4.6.2 Cell Culture

4.6.2.1 PBMC Isolation

To isolate the peripheral blood mononuclear cells (PBMCs), four Vacuette Heparin tubes (Greiner Bio-One, Kremsmuenster, Austria) were filled with 10 ml blood each. Two of the blood tubes were transferred into one 50 ml Falcon tube and filled to 35 ml with Roswell Park Memorial Institute (RPMI)-1640 (Sigma-Aldrich, Missouri, USA) (two per person). In a separate 50 ml Falcon tube, 15 ml of Histopaque (Sigma-Aldrich) was added (two per person). By holding the tube containing the Histopaque at an angle the blood/RPMI-1640 mixture was gently pipetted over the Histopaque, so as to create two layers. These 50 ml tubes were centrifuged at 400 x g for 25 minutes (no acceleration or brakes). This separated the tube contents into four layers. The top layer containing serum was discarded. The second thin layer contained the PBMCs, which were transferred to a new 50 ml Falcon tube and filled to 50 ml with RPMI-1640 (Sigma-Aldrich). Tubes were centrifuged at 250 x g for 5 minutes, the supernatant was decanted and cells were resuspended in 10 ml RPMI-1640 (Sigma-Aldrich). After centrifuging at 250 x g for 5 minutes, the supernatant was removed and the cells resuspended in 4 ml complete serum. Cells were plated in 6-well plates (NUNC, Roskilde, Denmark) with each well containing 2 ml resuspended cells and monocytes were allowed to adhere to the surface for 16 hours in an incubator at 37°C and 5% CO₂. Two wells were used per individual; one well to be subjected to BCG infection while the other would serve as a control

All experiments were conducted in triplicate.

4.6.2.2 BCG Infection

BCG was passed fifteen times through a 1ml/cc insulin syringe (Supra Latex, Cape Town, SA) without creating any bubbles. BCG stock solution was diluted by adding 800 µl RPMI-1640 (Sigma-Aldrich). After 16 hours, the media was removed and wells were washed once with 1 ml RPMI-1640 (Sigma-Aldrich) to remove all non-adherent cells. To each BCG well, 40 µl of diluted BCG was added and allowed to incubate at 37°C and 5% CO₂ for 4 hours. To each control well, 40 µl of RPMI-1640 (Sigma-Aldrich) was added.

After 4 hours, all media was removed and wells were washed once with 1 ml RPMI-1640 (Sigma-Aldrich). After the washing step, 2 ml of 10% conditioned media was added to each well. Plates were placed back into the incubator and 3 days post-infection a further 1 ml of 10% conditioned media was added. After 7 days the cells were lysed and RNA and proteins were extracted (*sections 4.6.3.1 and 4.6.4.1 respectively*).

BCG culture aliquots were a gift from Leanie Kleynhans (Immunology Group).

4.6.3 mRNA Expression

4.6.3.1 RNA Extraction

RNA was extracted from the monocytes using the TRIzol[®] Reagent (Invitrogen, California, USA). Briefly, plates were placed on ice and all media was removed. Each well was then washed once with 1 ml RPMI-1640 (Sigma-Aldrich). Working quickly, 500 µl of TRIzol[®] was added to the well and using a scraper the cells were lysed. The cell lysate was transferred to a chilled 1.5 ml tube. This was repeated with a further 500 µl of TRIzol[®]. The cell lysate solution was then allowed to incubate for 5 minutes at room temperature. To allow for phase separation, 200 µl of chloroform was added to each 1.5 ml tube and shaken vigorously for 15 seconds and then allowed to incubate for 3 minutes at room temperature. Samples were then centrifuged at 12 000 x g for 15 minutes at 4°C. After centrifugation, the solution separated into three phases, namely a lower red phenol-chloroform phase (containing proteins), an interphase (containing DNA) and a colourless upper aqueous phase (containing RNA). The aqueous phase was transferred to a new 1.5 ml tube containing 500 µl of isopropanol and samples were allowed to incubate for 15 minutes at room temperature. The organic phase was stored at -80°C for protein extraction at a later time (see *section 4.6.4.1*). Samples were centrifuged at 12 000 x g for 10 minutes at 4°C. The RNA formed a pellet at the bottom of the tube. The supernatant was discarded and add 1 ml 70% ethanol was used to wash the pellet. Samples were vortexed and then centrifuged at 7 500 x g for 5 minutes at 4°C. Ethanol was removed and RNA pellet allowed to air dry for 10 minutes (do not allow pellet to dry completely as this will reduce its solubility). RNA pellet was dissolved in 20 µl Diethyl Pyrocarbonate (DEPC) water (Ambion, Texas, USA) by pipetting or incubating for 10 minutes at 55°C to 60°C. RNA samples were stored at -80°C.

To determine the quantity and quality of RNA, we used the NanoDrop[®] ND-1000 Spectrophotometer and the NanoDrop[®] v.3.0.1 Software (Inqaba Biotechnology). An A260/A280 ratio < 1.6 meant that the RNA was partially dissolved.

4.6.3.2 Reverse-Transcription PCR (RT-PCR)

To synthesize cDNA, RT-PCR was done using the Transcriptor First Strand cDNA Synthesis Kit (Roche, Mannheim, Germany), following manufacturer's instructions. Briefly, in a 20 µl reaction

volume, 100ng of total RNA; 2 µl of primers (Random Hexamer Primer, 60 µM) and PCR-grade water (to make total volume equal 13 µl) was added to a nuclease-free PCR tube and heated for 10 minutes at 65°C to denature the RNA-primer mixture. The following reagents were then added in order: 4 µl 5X Transcriptor Reverse Transcriptase Reaction Buffer, 0.5 µl Protector RNase Inhibitor, 2 µl 10 mM dNTPs and 0.5 µl Transcriptor Reverse Transcriptase. Tubes were placed in a thermal block cycler with the following cycling conditions: 10 minutes at 25°C, 30 minutes at 55°C, 5 minutes at 85°C and then cooled down to 4°C to stop the reaction. cDNA samples were stored at -20°C.

4.6.3.3 Quantitative Real-Time PCR (qPCR)

To conduct the qPCR experiments the LightCycler® FastStart DNA Master SYBR Green I kit (Roche) was used with the LightCycler® 1.5 System (Roche). Firstly, the Master Mix was prepared by adding 10 µl of LightCycler® FastStart Enzyme to 54 µl 10X LightCycler® FastStart Reaction Mix SYBR Green 1. The PCR mix was prepared by adding, in order, the following components: 13.2 µl PCR-grade water, 0.8 µl MgCl₂, 3 µM of each primer, 2 µl Master Mix and 2 µl cDNA and pipetted into a precooled LightCycler® capillary. The capillaries were then centrifuged for 5 seconds at 700 x g. The qPCR program was as follows: Activation, 95°C for 15 minutes (ramp rate (R/R) of 20°C); PCR, 40 cycles of 95°C for 15 seconds (R/R of 20°C), 63°C for 15 seconds (R/R of 20°C) and 72°C for 20 seconds (R/R of 20°C) with a single acquisition; Melt Curve, 95°C for 0 seconds (R/R of 20°C), 60°C for 15 seconds (R/R of 20°C) and 95°C for 0 seconds (R/R of 0.1°C) with continuous acquisition; Cooling, 40°C for 30 seconds (R/R of 20°C). For each run a negative (no cDNA) control was included. For primers see Table 12.

Table 12: Primer sets used for qPCR analysis.

Type	Name	Sequence (5' – 3')	Size (bp)
Gene of interest	ANXA11F [§]	AGTGCCGAGCTACCCAGGATACC	403
	ANXA11R [§]	AGCGAGGATCTCAATCAGGCAGG	
Reference gene	TubulinF [*]	GCCAGAGCCAAGTGAC	60
	TubulinR [*]	CGAAGAAGGTGTTGAAGGAAT	

[§]Primers from Hoffman et al 2008

^{*}Primers from Roche

To determine the change in gene expression the Relative Quantification method was employed. This method measures the difference in the threshold cycles (C_q) between the gene of interest and a reference gene index and is expressed as the fold difference between a test (BCG infected) and a control (BCG uninfected) sample. To calculate the ratio of gene expression change (R) the following formulae were used:

$$\Delta C_q = C_q \text{ target} - C_q \text{ reference} \dots (1)$$

$$\Delta \Delta C_q = \Delta C_q \text{ sample} - \Delta C_q \text{ control} \dots (2)$$

$$R = 2^{-\Delta \Delta C_q} \dots (3)$$

Using equation (1) the difference in C_q was determined for each sample. Then by using equation (2) we determined the difference in C_q for each individual by comparing the C_q of the treated and control sample. Finally, R was calculated using equation (3), which is representative of the difference in the “corrected” number of cycles to threshold. The value of 2 is used based on the assumption that the product doubles with each cycle.

All analysis was done in duplicate.

4.6.4 Protein Expression

4.6.4.1 Protein Extraction

To the red-phenol chloroform phase (see *section 4.6.3.1*) 1 ml 100% ethanol was added and samples mixed by inversion followed by incubation at room temperature for 2 minutes. Samples were centrifuged at 2000 x g for 5 minutes, all centrifugation steps at 4°C. Phenol-ethanol supernatant was removed and 3 ml acetone was added. Samples were mixed by inversion for 10 – 15 seconds to obtain a homogenous solution. Samples were incubated for 10 minutes at room temperature and then the proteins were sedimented by centrifugation at 3000 x g for 5 minutes. Supernatant was discarded and pellet dispersed in 1 ml 0.3M guanidine thiocyanide in 95% ethanol solution. Incubation and centrifugation step was repeated. Supernatant was discarded and pellet washed with 1 ml 100% ethanol containing 2.5% glycerol (v/v) solution. Incubation and centrifugation step was repeated. Ethanol solution was decanted and pellet allowed to dry by inversion for 20 minutes. Pellet was resuspended in 150 µl 1% sodium dodecyl sulphate (SDS).

4.6.4.2 Standard Curve

To determine the protein concentration a standard curve was created using the D_c Protein Assay (Bio-Rad, Johannesburg, SA). Briefly, a Bovine Serum Albumin (BSA) series dilution (standards) was created (0 mg/ml BSA – 1 mg/ml BSA). Since our protein was resuspended in 1% SDS, a detergent, 20 µl of reagent S was added to each millilitre of reagent A used and relabelled reagent A'. 50 µl of standards and proteins (10-fold dilution) were pipetted into clean 15 ml Falcon tubes. 250 µl of reagent A' was then added to each tube and vortexed. 2 ml of reagent B was added to each tube and vortexed immediately. After 15 minutes the absorbance was read on a spectrophotometer (Ultrospec 4051, LKB Biochrom, Cambridge, UK) at 750 nm. The signal was stable for 1 hour.

The absorbance values of the standards were inserted into an Excel spreadsheet to generate a standard curve, with its associated equation ($y = mx + c$). This equation was used to calculate the concentration of each protein sample based on its absorbance value.

4.6.4.3 Western Blot

After preparing a 3% stacking and 12% separating sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) gel, 10 µl of PageRuler™ Prestained Protein ladder (Fermentas, Ontario, Canada) and 10 µl of sample (5 µg of protein plus 5 µl of reducing buffer (RB)) was loaded into each well and allowed to run for 2 hours at 200V in running buffer. After completion of electrophoresis of the gel, the Western Blot (WB) apparatus was set-up. Briefly (per gel), two pieces of Western Blotting Filter Paper (Thermo Scientific, Illinois, USA) was placed in transfer buffer. Hybond-P PVDF (GE Healthcare, New Jersey, USA) membrane was activated in 100% methanol, then washed for 5 minutes in distilled water and finally allowed to equilibrate for 10 minutes in Transfer Buffer. The WB sandwich was stacked as follows (black to red of WB cassette): filter paper, gel, membrane, filter paper. All air bubbles were removed and the sandwich was placed in WB Cell System. The WB Cell System was run under cold conditions (4°C) for 1 hour at 100V and placed on a magnetic stirrer to ensure distribution of cold transfer buffer throughout the run. After successful transfer of proteins from gel to membrane, the membrane was blocked overnight at 4°C in blocking buffer. After overnight blocking, the membrane was rinsed in TBS-T

buffer and then incubated with the primary antibody (Rabbit polyclonal antibody to Annexin X1, GeneTex[®], Inc., California, USA) (1:3000) for 1 hour at room temperature. Following incubation, the membrane was rinsed twice in TBT-T Buffer followed by washing three times, 10 minutes each, in TBS-T Buffer with shaking. The membrane was then incubated with the secondary antibody (Goat anti-rabbit IgG-HRP, Santa Cruz Biotechnology Inc., California, USA) (1:5000) for 1 hour at room temperature. The rinse/wash step was then repeated. To view the proteins on an autoradiographic film, the membrane was incubated in ECL Western Blotting Detection Fluid (GE Healthcare), following the manufacturer's instructions, for 1 minute. The membrane was then placed between two sheets of plastic and transferred to a Hypercassette[™] (Amersham Life Sciences, Buckinghamshire, England) with autoradiographic film for 10 minutes. The autoradiographic film was then developed using the Hyper Processor (Amersham Pharmacia Biotech, New Jersey, USA).

4.7 Methylation Pattern Analysis

4.7.1 Sample Selection

For the methylation experiments, eight individuals (four cases and four controls) were selected from the Case-Control sample set. Cases were selected based on the extent of their disease (number of cavities and size of cavities). Controls matching the cases in terms of age and gender were included.

4.7.2 Bisulfite Conversion of DNA

To determine whether or not the methylation pattern of ANXA11 differs between cases and controls the Imprint[™] DNA Modification Kit (Sigma-Aldrich) was used. This kit allows for DNA modification such that methylated cytosines are converted to uracil and unmethylated cytosines remain unchanged. Briefly, 1.1 ml of DNA Modification Solution was added to 1 vial of DNA Modification Powder and vortexed until the solution became clear. Once the solution was clear 40 μ l of Balance Solution was added (Solution A), followed by vortexing. To a clean 1.5 ml micro-centrifuge tube, 500 ng of DNA was added to 110 μ l of Solution A and briefly vortexed. Tubes were then incubated at 99°C for 6 minutes immediately followed by incubation at 65°C for 90 minutes.

Following the incubation, samples were placed on ice and subjected to a Post Modification DNA Clean-Up. For each sample a Spin Column was inserted into a Capless Collection Tube, to which 300 μ l Capture Solution was added and allowed to incubate for 1 minute. To this the modified DNA solution was added. The columns were then centrifuged at 12 000 x g for 20 seconds (all centrifugations were at 12 000 x g) and the flow-through discarded. 200 μ l of Cleaning Solution was then added to each column and centrifuged for 20 seconds. 50 μ l of Balance/Ethanol Solution was added to the bottom of the column and allowed to incubate for 8 minutes at room temperature. Columns were then centrifuged for 20 seconds. Then 200 μ l of 90% Ethanol Solution was added and columns were centrifuged for 20 seconds. The previous step was repeated but the columns were centrifuged for 40 seconds, the capless collection tube was discarded and the spin column placed in a clean 1.5 ml Collection Tube. 20 μ l of Elution Solution was added to the bottom of the spin column and allowed to incubate for 1 minute followed by centrifugation for 20 seconds. The spin column was discarded and the modified DNA stored at -80°C. The modified DNA was then ready for downstream applications.

4.7.3 Cloning of Bisulfite Converted DNA

4.7.3.1 Amplification and Purification of Bisulfite Converted DNA Fragments

The CpG island in the promoter region of *ANXA11* was amplified using two sets of primers (Table 13). PCR reactions were carried out in a 25 µl reaction containing: 2.5 µl 10X Buffer (containing 1.5mM MgCl₂) (Southern Cross Biotechnologies), 2 µl 1.25mM dNTPs (Bioline), 100 µM of each primer, 16.4 µl distilled water, 0.1 µl SuperTherm GOLD HotStart *Taq* polymerase and 2 µl of bisulfite converted DNA. The following cycling program was used: 10 minutes at 95°C, followed by 40 cycles of 20 seconds at 94°C, 30 seconds at 55°C and 40 seconds at 72°C, followed by an extension step of 2 minutes at 72°C and a cooling step of 2 minutes at 4°C. PCR products were visualized on a 2% agarose gel.

Table 13: Primers used for amplification of CpG island in promoter region of *ANXA11*.

Name	Sequence (5' – 3')	Size (bp)
Fp1	GAGGAAAGTTTTGAAGATAG	180
Rp1	CTCAATCTAACCTAAATAAAACC	
Fp2	GTTTTATTTAGGTTTAGATTG	195
Rp2	ACTTCCTAATTACTACTACAAA	

The PCR products were purified by centrifugation using the Wizard SV Gel and PCR Clean-Up System (Promega, Wisconsin, USA) following the manufacturer's instructions. Briefly, fragments of interest were cut out of the gel under UV light and placed in a clean 1.5 ml centrifuge tube. Membrane binding solution, the amount of which was calculated by determining the weight (grams) of the gel piece containing the fragment and adding 10 µl per 10 mg, was added to each tube. These tubes were then placed in a heating block (55°C) until the gel fragment was completely dissolved. The solution was then transferred to a minicolumn and allowed to incubate for 1 minute at room temperature. The minicolumns were then centrifuged at 14 000 rpm for 1 minute (all subsequent centrifugations are at 14 000 rpm). The flow-through was discarded and the minicolumn re-assembled. The minicolumns were then washed with 700 µl of Membrane Wash Solution and then centrifuged for 1 minute. After discarding of flow-through and re-assembling of minicolumns the wash step was repeated with 500 µl of Membrane Wash Solution and centrifuged for 5 minutes. Columns were then transferred to clean 1.5 ml centrifuge tubes, to which 50 µl of nuclease-free water was added to the centre of the membrane and allowed to incubate for 1 minute at room temperature. DNA was eluted from column by centrifugation for 1 minute. Samples were stored at -20°C.

4.7.3.2 Ligation to pGEM[®]-T Easy vector

The purified bisulfite converted DNA fragments were each ligated into the pGEM[®]-T Easy vector (Promega) using T4 DNA ligase (Promega). The principle behind the pGEM[®]-T Easy vector system is that the addition of adenosine bases to the 3' end of amplified fragments facilitates its ligation to the 3' tyrosine overhangs of the pGEM[®]-T Easy vector. This is made possible by the use of SuperTherm GOLD HotStart *Taq* polymerase which adds the adenosine bases to the 3' ends of amplified fragments. The ligation reaction was done in a 10 µl reaction containing: 5 µl 2X ligation buffer, 1 µl pGEM[®]-T Easy vector, 3 µl purified PCR product and 1 µl T4 DNA ligase. Ligation reactions were incubated overnight at 4°C.

4.7.3.3 *E.coli* Transformation

Competent *E. coli* JM109 (Promega) cells were transformed by means of heat shock. In summary, 25 µl of competent *E. coli* was added to 2.5 µl ligation product and allowed to incubate on ice for 10 minutes. Cells were then heat treated at 42°C for 47 seconds and then placed on ice for a further 2 minutes. 900 µl of Luria-Bertani (LB) medium (4°C) was added to the transformed cells and cells were placed in a shaker (Orbital Shaker Incubator LM-530, Yih Der, Japan) for 60 minutes at 37°C. Following incubation, cells were plated onto LB agar plates supplemented with 100 µg/ml isopropyl β-D-1-thiogalactopyranoside (IPTG), 100 µg/ml 5-bromo-4chloro-3-indolyl β-galactopyranoside (X-Gal) and 50 µg/ml Ampicillin (*E. Coli* FastGrow™ Media, Fermentas) and allowed to grow overnight at 37°C.

4.7.3.4 Colony PCR

Due to the addition of IPTG and X-gal to the growth medium, blue and white colonies were visible on the plates after incubation. White colonies were indicative of cells containing the plasmid and were therefore selected for further investigation. Using a pipette tip, a single white colony was selected and transferred to 4 ml LB broth containing 4 µl of ampicillin. These cell cultures were then placed in the orbital shaker for 8 hours at 37°C. After the incubation, a colony PCR was performed to determine if the cells contained the fragment of interest. For PCR conditions see Section 4.7.3.1. 2 µl of cell culture was substituted for 2 µl of bisulphite converted DNA.

4.7.3.5 Small-Scale Plasmid Extraction (Miniprep)

After confirmation that the cell cultures contained the correct size insert, plasmids were extracted using the Wizard Plus SV Miniprep Kit (Promega). Briefly, the *E. coli* cell cultures were pelleted by centrifugation for 10 minutes at 4 000 rpm and the supernatant discarded. Cells were then resuspended in 300 µl Cell Resuspension Solution by vigorous pipetting. Cells were then lysed by adding 350 µl of Cell Lysis Solution and tubes were inverted four times to mix solution. 10 µl of Alkaline Protease Solution was added which inhibits all endonucleases and other proteins released during lysis of cells, tubes were inverted four times and allowed to incubate at room temperature for 5 minutes. Following incubation, 400 µl of Neutralization Solution was added. After inverting the tubes four times they were centrifuged at maximum speed for 10 minutes (all subsequent centrifugations were at maximum speed). After centrifugation, the clear supernatant was transferred to a minicolumn and centrifuged for 1 minute. The column was then washed with 750 µl of Membrane Wash Solution and centrifuged for 1 minute. The wash step was repeated but this time with 250 µl of Membrane Wash Solution and centrifuged for 2 minutes. The plasmid was then eluted from the column by adding 50 µl of nuclease-free water and centrifuged for 1 minute. Plasmid DNA concentration was determined using the NanoDrop® ND-1000 Spectrophotometer and the NanoDrop® v3.0.1 Software (Inqaba Biotechnology). Samples were stored at -20°C.

4.7.4 Bisulfite Sequencing (Analysis)

To determine whether or not there was a difference in methylation pattern of ANXA11 between TB cases and controls, the purified plasmids containing the bisulphite treated DNA were sent for sequencing. 5 µl of 100 ng/µl DNA was sent to the Central Analytical Facility (Stellenbosch University) for sequencing using the universal primer SP6 (provided by the Central Analytical Facility). Sequencing data was analysed using the BiQ Analyzer software package ([43](http://biq-</p></div><div data-bbox=)

analyzer.bioinf.mpi-inf.mpg.de/), which allows for visualization and quality control of DNA methylation data from bisulfite sequencing.

4.8 Statistical Analysis

4.8.1 Hardy-Weinberg Equilibrium

The Hardy-Weinberg model is used to determine the genotypic frequencies in a population by using the allele frequencies^{248, 249}. This model however is only valid in stable populations, that is, a population in which no gene flow, genetic drift, mutation, natural selection or non-random mating is present²⁵⁰. When all these assumptions hold true in a population and the population is of a large size, the genotypic and allelic frequencies will remain constant from generation to generation, resulting in a population that is in Hardy-Weinberg Equilibrium (HWE). HWE is therefore an important statistical tool in case-control association studies, since departure from HWE in the control study participants could mean that there was a genotyping error²⁵¹ or that the population being investigated is stratified; or in the event that the disease cases are not in HWE, the marker is associated with disease outcome²⁵².

The HWE model has two equations, $p + q = 1$ for allele frequencies and $1 = p^2 + 2pq + q^2$ for genotype frequencies, where the frequency of the one allele is p and the other is q . When one or both allelic frequencies are known, they can be used to determine the expected genotypic frequencies of the population, where p^2 and q^2 are the proportion of homozygotes and $2pq$ is the proportion of heterozygotes. These expected genotypic frequencies are then used to determine whether or not the population is in HWE by using the chi-square test (X^2 , 5.9.2). If the X^2 value is less than 3.84 then the population is considered to be in HWE. HWE analysis has one degree of freedom (df, number of genotypes minus number of alleles).

4.8.2 Chi-square Test

To test for HWE, a chi-square (X^2) goodness-of-fit test is employed^{251, 252}. The X^2 test is most commonly used to test if there is no association between two variables. The X^2 value is calculated using the following equation:

$$X^2 = \sum \frac{(\text{observed number} - \text{expected number})^2}{\text{expected number}}$$

Where the expected number is the number of individuals with that genotype expected if the population is in HWE and the observed number is the number of individuals obtained from the data. When the observed and expected values are similar then the hypothesis of no association is held true, while if the observed and expected values differ significantly then there is an association between the two variables. The P -value, which is calculated from the X^2 test and its degrees of freedom, is used to determine if the null hypothesis (H_0 : there is no significant difference between the observed and expected frequencies) or the alternative hypothesis (H_1 : there is a significant difference between the observed and expected frequencies) is true. If $P < 0.05$ then the H_0 is rejected and the H_1 is accepted.

The X^2 value was determined using Prism v5.02 (GraphPad Software, Inc., California, USA).

4.8.3 Fisher's Exact Test

Fisher's exact test was done using Prism v5.02. This test is used to analyse 2 x 2 contingency tables with small sample sizes²⁵³, and was used to calculate an exact *P*-value, whereas the X^2 test is an approximation.

4.8.4 Haplotype and Linkage Disequilibrium

Haplotype analysis was done using Cocophase in the Unphased suite²⁵⁴ and 10 000 permutation replicates were done to test for global significance of the estimated haplotypes. Since haplotypes in a haplotype block are not independent of one another the use of Bonferroni corrections would be too conservative, as it would result in haplotypes with moderate effect not being investigated. LD was evaluated using Haploview v4.1²⁵⁵, with haplotype blocks being automatically selected by the programme, based on the block definition of Gabriel *et al*¹⁰⁹. LD was analysed in terms of D' , where $D' = 0$ indicates no LD and $D' = 1$ equals complete LD.

4.8.5 Power Calculations

Power calculations were done using Epi Info 2000 (Centres for Disease Control and Prevention, USA). With the number of samples available we had 80% power and 95% confidence given an expected allele frequency of at least 5% for the SNPs genotyped to detect an odds ratio of 2.23.

Chapter 5: ANXA11

Perfect as the wing of a bird may be, it will never enable the bird to fly if unsupported by the air. Facts are the air of science. Without them a man of science can never rise.

Ivan Pavlov

5.1 Annexins

5.1.1 Annexin Gene Family

The annexins are a superfamily of calcium-dependent phospholipid binding proteins that are structurally related²⁵⁶⁻²⁵⁸. There are currently five classes of annexins (A – E), with human annexins belonging to class A, comprised of 12 annexins (A1 – A11 and A13)^{257, 258}. All annexins contain a core domain which consists of a 70 amino acid repeat, which occurs either four (35-37 kDa) or eight (~65 kDa) times^{256, 259, 260}, and is called the annexin repeat. Each annexin repeat is made up of five α helices and usually contains a 'type 2' motif²⁵⁸, which is responsible for the binding properties of annexins to calcium (Ca^{2+}) and phospholipids^{257-259, 261, 262}. These four or eight annexin repeats make up the C-terminal core domain and are highly conserved in all annexin proteins. Conversely, the N-terminal domain varies greatly between the annexin proteins, and it is this variability that gives rise to the functional differences in biological activities of each annexin protein^{257, 260, 262}.

Annexin proteins are generally found within the cytosol, while under certain conditions some annexins (A2 and A11) have been found to localize in the nucleus or be expressed on the cell surface (A1 and A2) even though no secretory signal peptide has yet been identified²⁵⁸. Annexin A1 (ANXA1) has been found to translocate to the cell surface after cell exposure to glucocorticoids, while annexin A2 (ANXA2) is constitutively expressed on vascular endothelial cells where it plays a role in the regulation of blood clotting²⁵⁸ and maintenance of fibrinolytic homeostasis²⁶¹. Annexin protein expression levels and tissue distribution has been found to vary between family members, with some annexins being abundant and ubiquitously expressed (annexins A1, A2, A4, A5, A6, A7 and A11), while others are selectively expressed (annexin A3 – neutrophils and annexin A8 – placenta and skin) or restrictively expressed (annexin A9 – tongue, annexin A10 – stomach and annexin A13 – small intestine)²⁵⁸.

Due to the ubiquitous expression of most of the annexin proteins, researchers have postulated that these proteins must play a fundamental role in various cellular biological processes²⁶³ such as apoptosis, cell division, cell signalling, endo- and exocytosis and ion transport^{257, 261, 262}. With the use of mouse KO models, various functional insights have been identified for some annexins. For example, KO of *ANXA1* resulted in an altered inflammatory response and glucocorticoid effects²⁶⁴ while *ANXA7* null mutant mice strains were shown to be either embryonic lethal²⁶⁵ or have an altered Ca^{2+} homeostasis²⁶⁶. However, phenotypic changes noted in cell culture studies are not always seen in mice KO models, or are much more subtle, highlighting the potential functional redundancy of annexin proteins^{258, 262}. The functional role of each annexin protein has not yet been identified and further studies are required.

5.1.2 Annexin A11

Annexin A11 (ANXA11), is located on chromosome 10q22.3 and is considered to be the most evolutionary ancient member of the annexin superfamily of proteins²⁶⁷. It arose from the duplication event of annexin A13 into *ANXA7* and *ANXA11*²⁶⁸. *ANXA11* comprises a C-terminal core that is shared between all annexin genes and contains the Ca^{2+} binding domain^{268, 269} and a unique long N-terminal domain rich in glycine, proline and tyrosine residues, which has been shown to contain the binding sites for the apoptotic proteins calcylin (S100A6) and the apoptosis-linked protein (apoptosis-linked gene-2, ALG-2)^{267, 269}, of which binding has been shown to be

calcium dependent in *in vitro* studies²⁶⁷. In addition, it is believed that this N-terminal region is responsible for the protein's autoantigenicity, nuclear localization, extracellular targeting and tyrosine phosphorylation properties^{268, 270}. Previous studies have also shown a possible role for ANXA11 in apoptosis, calcium signalling, cell proliferation, insulin secretion and vesicle trafficking²⁶⁷⁻²⁶⁹. Expression of ANXA11 is ubiquitous with localization predominantly in the nucleus²⁷¹.

5.1.3 Disease Mechanisms – ANXA11

5.1.3.1 Cancers

Annexins have been studied in various cancers²⁷²⁻²⁷⁵, and ANXA11 plays an important role in ovarian^{269, 276} and colorectal²⁵⁷ cancer. In a study conducted in 2007 by Song *et al.* in ovarian cancer patients, ANXA11 was found to be associated with the development of cisplatin (chemotherapeutic drug) resistance and related to tumour recurrence. It was found that in three cisplatin-resistant cell lines ANXA11 was continuously down-regulated when compared to the parental cells. They also observed normal ANXA11 expression in most “normal” human organs while decreased levels of ANXA11 was found in various human malignancies. Furthermore, they found that increased ANXA11 protein levels resulted in ovarian cancer patients being disease-free for longer time periods. In a 2009 follow-up study, Song *et al.* conducted a study to elucidate the mechanism by which ANXA11 suppression resulted in cisplatin chemoresistance using small interfering RNA. They showed that by knocking down ANXA11 expression the ability of ovarian cancer cell lines to proliferate and form colonies was reduced, and therefore that epigenetic silencing of ANXA11 conferred cisplatin resistance on these cell lines.

In 2008, Duncan *et al.* characterised the role of annexins in colorectal cancer. By using comparative proteomic analysis they identified annexins A1, A2, A4 and A11 as being overexpressed in colorectal tumour samples. They found increased expression of ANXA11 in primary tumours compared to normal colon tissue, and increased ANXA11 expression correlated with an increase in tumour stage. They also noted that individuals who had higher annexin A4 and A11 expression had a lower chance of survival. They postulated that due to the role of ANXA11 in cell growth, dysregulation of ANXA11 could result in tumour progression. However, the mechanism by which the gene is regulated has yet to be elucidated.

5.1.3.3 Sarcoidosis

In a recent GWAS for sarcoidosis in a German population, Hofmann *et al.* identified *ANXA11* as a novel susceptibility factor²⁷⁷. Sarcoidosis is a multisystem immune disease that mostly affects the lungs and lymphatic system^{278, 279} but is of unknown aetiology²⁸⁰. It is however believed that the causative agent could be of microbial origin and that it is airborne since the disease mostly affects the lungs^{281, 282}. As is the case with TB, sarcoidosis results in the formation of granulomas^{278, 279} in diseased individuals with symptoms which include weight loss, extensive coughing, fever, fatigue and night-sweats^{279, 283}. Various polymorphisms associated with *ANXA11* in the GWAS were found to alter disease outcome in sarcoidosis patients. The authors postulated that the reduction or presence of aberrant forms of ANXA11 could potentially alter the apoptosis pathway by destroying the balance between apoptosis and the survival of activated inflammatory cells in sarcoidosis patients²⁷⁷.

In the recent GWAS for novel susceptibility variants for sarcoidosis in the German population, *ANXA11* polymorphisms were found to alter susceptibility in the population²⁷⁷. Due to the similarity in the pathomechanisms between sarcoidosis and TB²⁸⁴, we were interested in

identifying whether or not these *ANXA11* polymorphisms could also alter susceptibility to TB in our SAC population. In the study by Hofmann *et al.*, six polymorphisms (Table 14) were found to be in strong LD with each other, in addition to being strongly associated with sarcoidosis individually.

Table 14: *ANXA11* polymorphisms investigated as TB susceptibility variants in the SAC population.

Polymorphism	Location
rs1049550	exon 6
rs2784773	intron 11
rs2573346	intron 12
rs1953600	3' UTR/downstream
rs7071579	3' UTR/downstream
rs2789679	3' UTR/downstream

5.2 Results

5.2.1 Genotype Analysis

All polymorphisms investigated by TaqMan® genotyping had a call rate $\geq 95\%$ and were in HWE. Bonferroni correction for multiple testing was done by determining the number of independent LD blocks²⁸⁵. Three LD blocks were determined for *ANXA11* and a P-value < 0.017 was adopted as a threshold for significance. Based on this cut-off value, the 3' UTR polymorphism rs7071579, was the only SNP to be significantly associated with susceptibility to TB (Table 15) with the A-allele being overrepresented in TB cases (OR 1.38, 95% CI 1.13-1.68). The intron 11 SNP (rs2784773) was found to be nominally significant with susceptibility to TB. The four remaining polymorphisms shared similar genotype and allele frequencies between cases and controls and were therefore not associated with susceptibility to TB in our population.

Table 15: Statistical analysis of *ANXA11* polymorphisms.

SNP	Allele		Genotype frequencies								P _{gen} [†]	P _{allele} [*]	OR [95%CI] [§]
			Cases				Controls						
	1 [‡]	2	11	12	22	HWE [#]	11	12	22	HWE			
rs2789679	T	A	0.06	0.33	0.61	0.40	0.06	0.35	0.59	0.63	0.94	0.81	1.03 [0.82-1.30]
rs7071579	A	G	0.33	0.45	0.21	0.11	0.25	0.45	0.29	0.080	0.01	0.0017	1.38 [1.13-1.68]
rs1953600	T	C	0.07	0.38	0.55	0.92	0.07	0.38	0.55	0.87	0.97	0.91	1.02 [0.82-1.27]
rs2573346	A	G	0.06	0.33	0.61	0.34	0.05	0.35	0.60	0.94	0.82	0.91	0.98 [0.78-1.24]
rs2784773	T	C	0.27	0.48	0.25	0.35	0.20	0.51	0.29	0.49	0.05	0.03	0.80 [0.66-0.98]
rs1049550	A	G	0.07	0.38	0.55	0.86	0.07	0.37	0.56	0.5	0.94	0.96	1.01 [0.81-1.26]

[‡]Minor allele in controls.

[†]P value from a genotype-based X²-test, significant if $P < 0.017$.

^{*}P value from an allele-based Fisher's exact test, significant if $P < 0.017$.

[#]Hardy-Weinberg Equilibrium.

[§]Odds Ratio and [95% Confidence Interval]

5.2.2 Linkage Disequilibrium and Haplotype Analysis

Haploview identified one haplotype block (Figure 11), which showed the intron 12 polymorphism (rs2573346) and the three 3' UTR polymorphisms (rs1953600, rs7071579 and rs2789679) to be in strong to moderate LD with each other, while the exon 6 (rs1049550) and intron 11 (rs2784773) polymorphisms were not in LD. The Cocaphase program identified three proposed haplotypes to be associated with susceptibility to TB in the SAC population (Table 16). Two of these haplotypes contained the A-allele of rs7071579 and were more frequent in cases than controls. The global P value, as calculated by Cocaphase, for the estimated haplotypes was also significant ($P = 0.0048$).

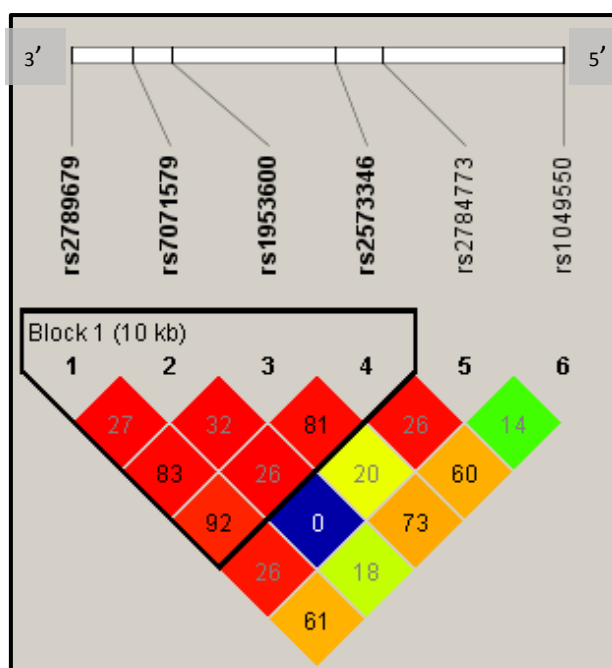


Figure 11: Plot of LD between the *ANXA11* markers in the SAC population, generated by Haploview v4.1. The 5' and 3' ends of the genes are indicated and r^2 values (%) are indicated on the squares. The colours of the squares represent D' values, with red being $D' = 1$ and blue $D' = 0$.

Table 16: Haplotype analysis for *ANXA11* polymorphisms.

Haplotypes*						Frequency		P value	OR
rs2789679	rs7071579	rs1953600	rs2573346	rs2784773	rs1049550	Cases	Controls		
A	A	C	G	C	G	0.024	0.01	0.048	3.06
A	G	C	G	C	G	0.23	0.29	0.01	1.05
A	G	C	G	T	A	0.028	0.016	0.1	2.3
A	A	C	G	T	G	0.28	0.20	0.00037	1.8
A	G	C	G	T	G	0.17	0.21	0.1	1.1
A	A	T	G	T	A	0.027	0.022	0.5	1.5
T	A	T	A	C	A	0.19	0.20	0.4	1.2
T	A	T	A	C	G	0.026	0.019	0.4	1.7
Global significance#								0.0048	

* Estimated haplotypes, as produced by Cocaphase.

Permutation test P value calculated from 10 000 permutations in Cocaphase, to correct for multiple testing while taking into account the correlation between markers and haplotypes.

5.2.3 Sequencing

Exons 4, 5, 6 and the 3' UTR region of *ANXA11* were amplified and the products visualized on a 1.5% agarose gel (Figure 12). The PCR products were then sequenced, and the resultant chromatograms were analysed using the Sequencher v4.7 software. Sequencing quality was determined by means of background signal levels, with low levels indicative of good quality sequencing.

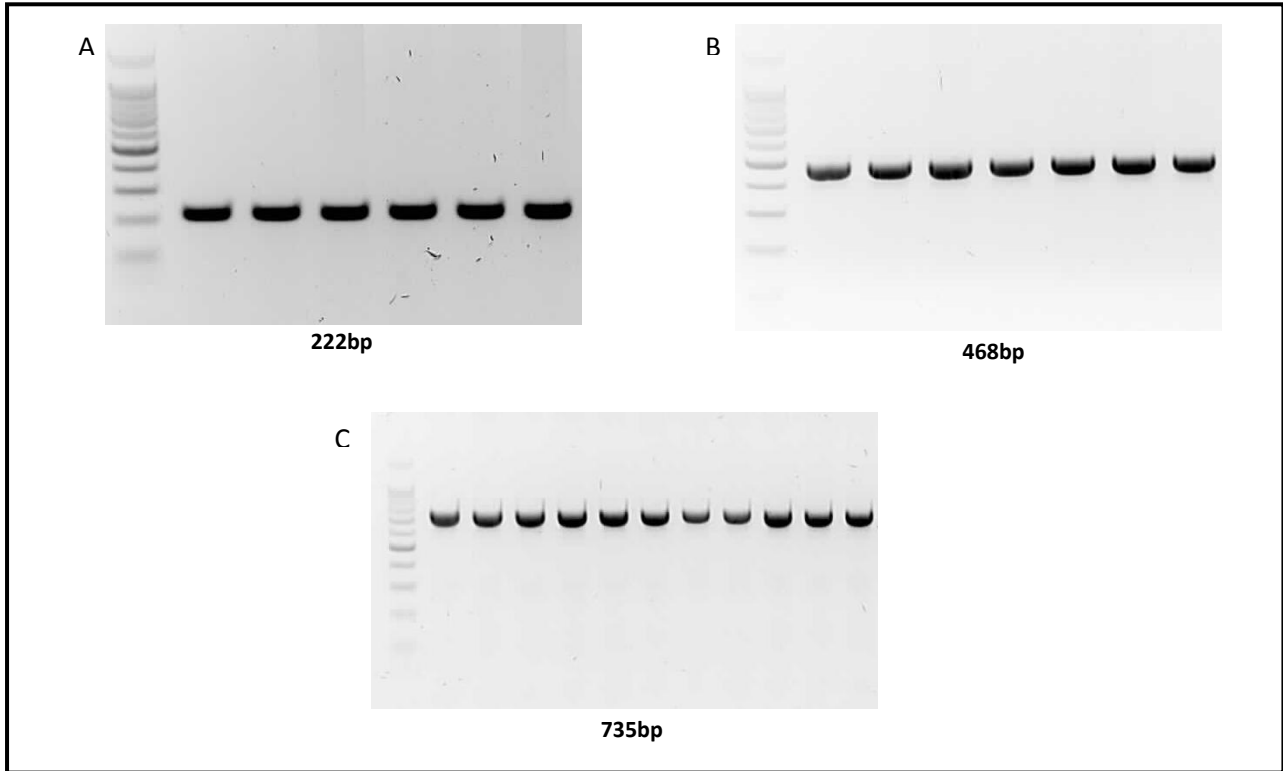


Figure 12: 1.5% agarose gel with amplified PCR products. A. Exon 4 amplified products, B. Exon 5 and 6 amplified products, C. 3' UTR amplified products. A 100bp DNA ladder was used, size of amplified products are indicated below each gel.

Using the Sequencher v4.7 software, the chromatograms of all the samples were aligned to each other and a reference sequence to identify any polymorphisms present (Figure 13). Two polymorphisms were identified in each of intron five and exon six of *ANXA11*, of which three were known and one was novel. For the 3' UTR region, seventeen polymorphisms were identified, with two previously characterized and fifteen novel. The sequencing quality of exon 9 was not good and could not be analysed. For allele frequencies, chromosomal positions and rs-numbers see table 17.

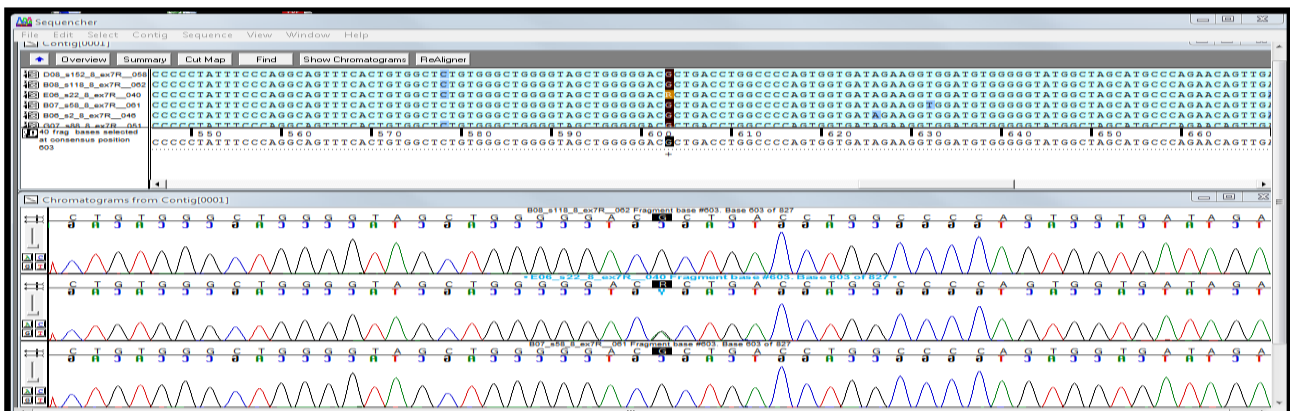


Figure 13: Alignment of sequence data using the Sequencher v4.7 software.

Table 17: Polymorphisms identified during sequencing of ANXA11 gene regions.

Location	Chromosomal Position [§]	rs-number	Frequency*
Intron 5	81, 916, 868	novel	0.05
Intron 5	81, 916, 730	rs34332933	0.2
Exon 6	81, 916, 698	rs2228427	0.2
Exon 6	81, 916, 682	rs1049550	0.3
3'UTR	81, 915, 479	novel	0.5
3'UTR	81, 915, 423	novel	0.55
3'UTR	81, 915, 406	novel	0.6
3'UTR	81, 915, 356	novel	0.6
3'UTR	81, 915, 304	novel	0.6
3'UTR	81, 915, 264	novel	0.6
3'UTR	81, 915, 260	novel	0.6
3'UTR	81, 915, 227	novel	0.6
3'UTR	81, 915, 205	novel	0.6
3'UTR	81, 915, 196	novel	0.6
3'UTR	81, 915, 178	novel	0.6
3'UTR	81, 915, 173	novel	0.05
3'UTR	81, 915, 147	novel	0.55
3'UTR	81, 915, 144	rs76487007	0.05
3'UTR	81, 915, 136	rs2789686	0.3
3'UTR	81, 915, 121	novel	0.6
3'UTR	81, 915, 117	novel	0.6

[§]Position relative to chromosome 10

*Frequency of individuals harbouring the polymorphism, n = 20

5.2.4 Gene Expression

The DNA sequence containing rs7071579 polymorphism was amplified in the DNA obtained from the volunteers and sent for sequencing. After analysing the sequencing data with Sequencher v4.7, we selected two individuals who were homozygous for the A-allele and one individual homozygous for the G-allele (the GG genotype is rare) for further investigation.

The ANXA11 expression levels with respect to the rs7071579 polymorphism, was analysed by qPCR and Western Blots (WB). To measure the difference in mRNA levels between individuals with the different genotypes, the cDNA transcribed from the RNA of the three individuals was analysed by means of qPCR. The ratio of gene expression change (R) for individuals homozygous for the A-allele was 1.28, whereas homozygosity for the G-allele resulted in R = 0.17 (Figure 14), where R = 1 is indicative of no change in gene expression levels. PCR efficiency equalled 0.98.

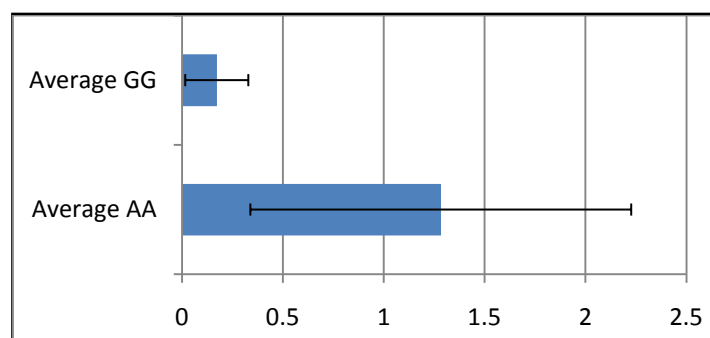


Figure 14: Differences in mRNA levels between individuals with different genotypes.

Unfortunately, no data was generated regarding the change in protein levels. After several rounds of optimization of the WB protocol, including antibody concentrations and incubation times in ECL detection fluid, we were still unable to generate a WB showing the ANXA11 protein. After consulting with experts in WB and a thorough review of our experimental technique, I think that the most likely reason for the failure to detect proteins on the WB was antibodies that were no longer active. Due to time constraints we were unable to order new antibodies and repeat the experiments. However, we do intend to revisit and complete this work in the future.

5.2.5 Methylation

To determine the role of methylation in TB susceptibility, we investigated the DNA methylation pattern of the promoter region of ANXA11 in eight individuals (four cases and four controls). After bisulfite treatment of the DNA, it was amplified (Figure 15), purified and ligated into the pGEM[®]-T Easy vector system. *E. coli* JM109 cells were heat shocked to facilitate the uptake of the vector containing the fragment of interest. The addition of IPTG and X-Gal to the growth medium on which the *E. coli* were cultured resulted in the formation of white and blue colonies (Figure 16), with white colonies containing the vector. To determine if the vector contained the insert, a colony PCR was conducted (Figure 17). Vectors were then purified and sent for sequencing.

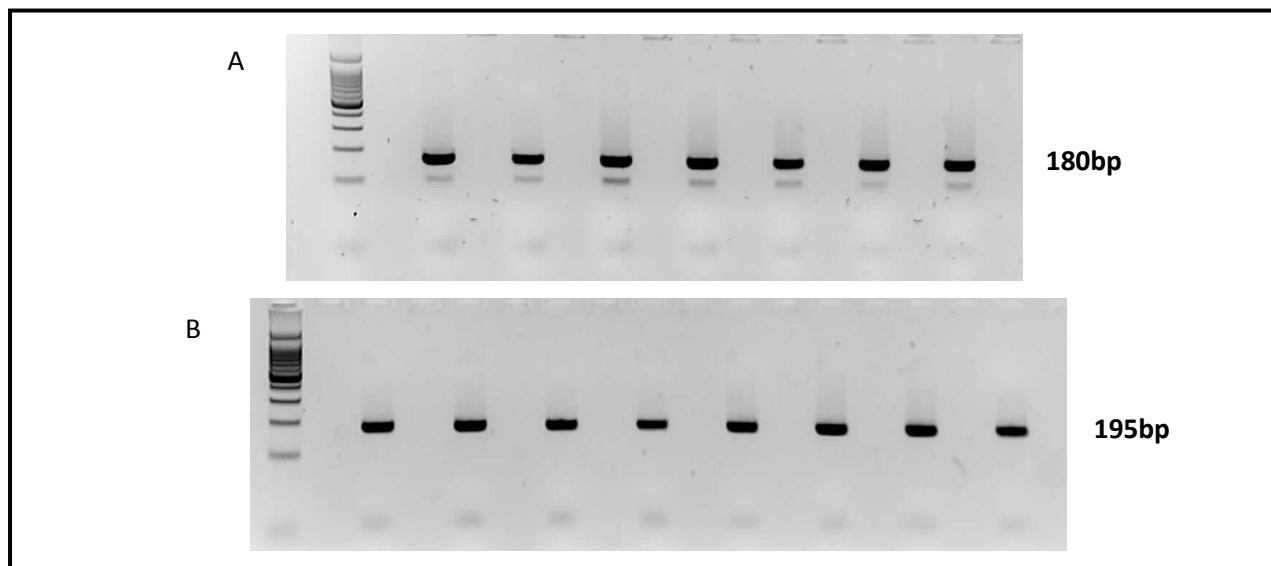


Figure 15: 2% agarose gel with amplified bisulfite treated PCR products of the promoter region of ANXA11. A. Promoter region amplified with primer set 1, B. Promoter region amplified with primer set 2. A 100bp ladder was loaded in the first lane. Fragment sizes are indicated.

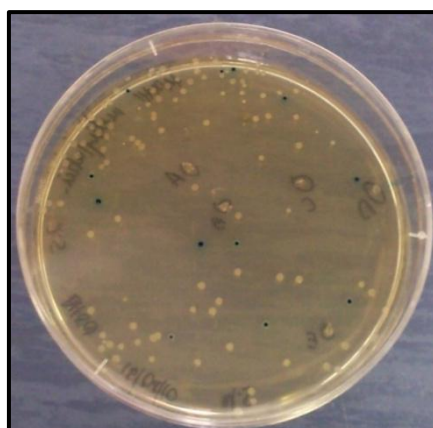


Figure 16: Blue/White colony selection of *E. coli* JM109 cells.

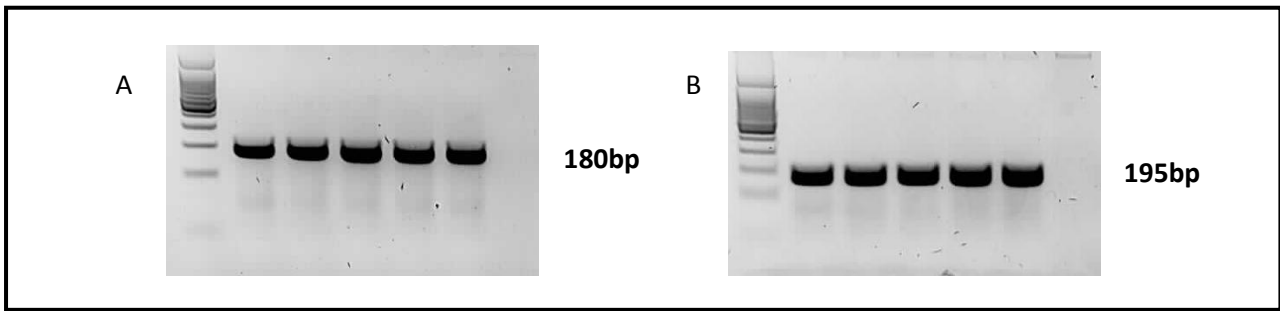


Figure 17: Colony PCR of *E. coli* JM109 cells for the verification of DNA fragment of interest. A. Colony PCR with primer set 1, B. Colony PCR with primer set 2. A 100bp DNA was loaded in the first lane. Fragment sizes are indicated.

Using BioEdit, the sequencing data was analysed, cleaned-up and aligned to the original DNA sequence and the “treated” DNA sequence and saved in the FASTA file format. The FASTA files were uploaded into the BiQ Analyzer Software package and methylation patterns of the CpG islands were interrogated, with no difference in methylation pattern of the ANXA11 promoter observed (Figure 18). Since all samples shared the same methylation pattern, only a subset of the data is shown in Figure 18.

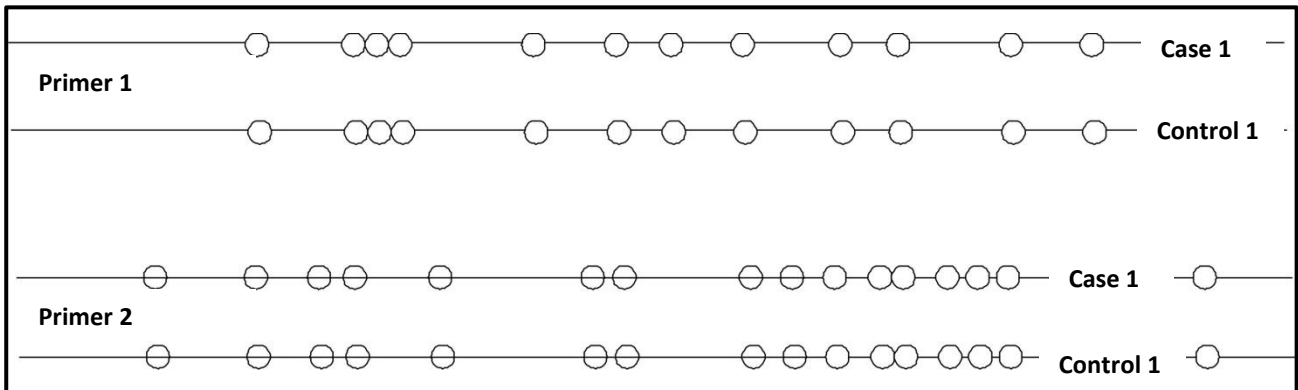


Figure 18: Lollipop diagram, with circles representative of CpGs, indicating no differential methylation between cases and controls. Empty circles means not methylated.

5.3 Discussion

In this study we identified a 3' UTR polymorphism of *ANXA11* that could alter susceptibility to TB, in our high incidence setting. *ANXA11* and its associated polymorphisms were first identified as a possible susceptibility factor for sarcoidosis in a GWAS in a German population²⁷⁷. In that study, the authors showed the rs1049550 polymorphism ($P = 1.0 \times 10^{-12}$, OR[95%CI] = 0.62[0.54 – 0.71]) to be strongly associated with sarcoidosis. The SNPs rs2784773, rs2573346, rs1953600, rs7071579 and rs2789679 were all in strong LD with this SNP ($r^2 = 0.44 – 0.98$). As sarcoidosis shares various similarities in pathophysiology to TB²⁸⁴, we decided to investigate these *ANXA11* polymorphisms as susceptibility factors for TB. In this study, we did not find the rs1049550 polymorphism to be a risk factor for TB. We did however observe the 3' UTR polymorphism rs7071579 to be associated with TB, with the A-allele occurring more frequently in cases than controls (56% vs 48%). In the three estimated haplotypes shown to be significantly associated with TB, the A-allele occurred in two of the haplotypes more frequently in cases, further suggesting that the A-allele increases the risk for developing TB. A comparison of allele frequencies between the German population and two African populations (South African Coloured and Yoruba in Ibadan, Nigeria) shows the A-allele

of rs7071579 to be the major allele in Germans (57%), whereas in the African populations the G-allele is the major allele (52% and 53% respectively) (Table 18).

Comparison of the LD between the South African and German populations reveals differences in the genetic structure of the two populations, with the German population having stronger LD (one LD block vs three LD blocks). This is contradictory to current opinions, since stronger LD²⁸⁶ would be expected in the SAC population as they are a highly admixed population^{237, 287}. This stronger LD may still be true over the whole genome. The differences in allele frequencies and LD structure discussed above indicate that susceptibility factors identified in the German population may not serve as disease factors in the South African population. In addition, it has to be kept in mind that the two diseases, although similar, do not necessarily share the same genetic risk factors.

Since an association was noted in the case-control association study, we believed that the ANXA11 polymorphisms warranted further experimental investigation into their role in TB susceptibility. The expression analysis conducted in this study showed that individuals who were homozygous for the A-allele of rs7071579 had approximately five times higher levels of ANXA11 mRNA than individuals homozygous for the G-allele. This finding contradicts the initial postulation of this study and that of Hofmann et al. (2008). It was postulated that individuals who have lower levels or aberrant copies of ANXA11 would be at greater risk of developing disease, with the case-control study that we conducted showing the A-allele to be associated with increased susceptibility to TB. Due to our limited understanding of the role of ANXA11 in infectious disease mechanisms, particularly apoptosis, how intracellular levels of ANXA11 would alter disease processes is not known. It could be that the mRNA levels identified are not a true reflection of protein levels intracellularly²⁸⁸, since other epigenetic factors could also play a role during translation. Lastly, the number of individuals included in this study was very small which therefore limits our confidence in the mRNA levels present. We are aware of the shortcomings of this study. Due to limited time, we were not able to address these and this study therefore serves as a pilot. We have increased the number of individuals of each genotype and decided to include individuals of heterozygous genotype and use two reference genes.

DNA methylation, the most well characterised epigenetic modification process, is the process whereby methyl groups are added to the cytosine of CpG sites and it is known to play a role in transcriptional regulation^{289, 290}. This occurs through its ability to change the expression of genes by influencing the binding affinities of transcription factors or other proteins to the promoter region of the gene²⁹¹. Hypermethylation of the promoter region is usually associated with no transcriptional activity due to the tightly bound nature of the chromatin, whereas hypomethylation is usually associated with lower levels of transcription when compared to unmethylated promoter regions. In this study we were interested in identifying whether the promoter region of ANXA11 was differentially methylated between cases and controls, and could explain the difference in expression levels, but no such difference was observed in the methylation patterns. This however is not surprising as our laboratory previously investigated 27 578 CpG sites across the entire genome using the Infinium HumanMethylation27, RevB BeadChip Kit from Illumina. With regards to ANXA11, two CpG sites were interrogated in four samples, with no difference in methylation pattern observed between cases and controls (unpublished). We were therefore interested in whether or not differences in global methylation of the ANXA11 promoter were present in this study. The role of DNA methylation in infectious diseases is still unknown. However, in other diseases, especially cancer, the role of methylation has been well characterized and has led to the development of novel treatments²⁹². Understanding how methylation affects TB disease outcomes is therefore worth investigating.

ANXA11 has been shown to play a role in various biological processes in the cell, including apoptosis, cell division and vesicle trafficking. In the recent GWAS conducted by Hofmann *et al.* (2008) they postulated that reduced or aberrant forms of ANXA11 could potentially alter the apoptosis pathway by destroying the balance between apoptosis and the survival of activated inflammatory cells in sarcoidosis patients. This theory could also be applicable to TB. Tomas *et al.* (2004) showed that daughter cells in which the *ANXA11* gene had been silenced using siRNA were unable to successfully separate from each other resulting in death by apoptosis. Recently, Song *et al.* (2009) knocked down ANXA11 expression in ovarian cancer cell lines, resulting in reduced cell proliferation and colony forming capability in these cells. To date, several studies have investigated the role of macrophage apoptosis in TB disease, with the role that it plays in the pathophysiology of the disease still not fully understood^{293, 294}. There are currently two hypotheses regarding the role of apoptosis in TB. The first is that macrophage apoptosis could in fact benefit the bacterium by inhibiting the host from eliciting an effective immune response due to the lack of an efficient phagosome and thus poor local inflammation²⁹⁵. The current model for this hypothesis is based on the fact that elimination of alveolar macrophages and resultant poor inflammation could result in a delayed or interrupted induction of the cell mediated immune response. However, the second hypothesis regarding the role of apoptosis in TB is that it is beneficial to the host with the overwhelming amount of scientific data in favour of this. Evidence includes the observation that virulent mycobacterial strains (H37Rv, Erdman) are much weaker inducers of macrophage apoptosis compared to attenuated strains of mycobacteria (H37Ra and BCG)²⁹⁶. This has been shown to be mediated through the ability of virulent strains to interfere with TNF- α signalling and upregulate the expression of Mcl-1 which has known anti-apoptotic properties²⁹⁷⁻²⁹⁹, thus highlighting the function of TNF- α mediated apoptosis in the host defensive response. Interestingly, certain viruses have been shown to employ similar mechanisms to inhibit pro-death signals of infected cells.

Much work has shown that induction of apoptosis results in a reduction in the viability of the mycobacteria whereas mycobacteria from macrophages undergoing necrosis exhibited no reduction in viability, illustrating that apoptosis exerts a direct antimicrobial effect on intracellular mycobacteria^{296, 300-303}. Other studies have shown that due to macrophage apoptosis a resultant altered interaction also occurs between the infected macrophages and other host defence cells. For example, when mycobacteria were presented to macrophages in apoptotic bodies, a reduction of bacterial growth of >90% was observed^{304, 305} and when presented to dendritic cells adaptive immunity was promoted²⁹⁴. Lastly, macrophage apoptosis results in less lung tissue destruction, and thus when it predominates over cell lysis, the lung architecture is preserved and spread of the bacterium to new hosts is prevented^{293, 306}. Therefore, knowing the importance of apoptosis in the outcome of diseases such as TB^{294, 307}, we could postulate that ANXA11 may play an important role in the risk of developing TB.

Another mechanism whereby ANXA11 could alter susceptibility to TB is via its potential role in the endocytic-autophagic pathway³⁰⁸⁻³¹⁰, given a number of findings which have demonstrated a role of ANXA1, 2 and 6 in the endocytic pathway^{259, 260}. One of the hallmark features of *M. tuberculosis* is its ability to prevent phagolysosome maturation in the infected macrophage, thus enabling it to avoid the key bactericidal mechanism of the invaded macrophage. This results in reduced antigen processing and presentation³¹¹.

Autophagy is the cellular process whereby the cell degrades its own components by lysosomal activity, and plays an important role in cell growth, development and homeostasis³¹². It has been

shown to be a vital survival process under starvation conditions. Recent studies have shown a role for autophagy in infectious diseases ³¹³⁻³¹⁵, especially against *M. tuberculosis* infection ^{311, 312}. Its role in the fight against invading pathogens is two-fold. Firstly, autophagy is able to facilitate antigen presentation and thus enhance the immune recognition of infected cells ³¹⁵. Secondly, although the mechanism is still unclear, it is able to kill intracellular mycobacteria. There are currently three potential models for how this could occur. These include: (i) the envelopment of the arrested mycobacterial phagosome by the autophagosome with subsequent autolysosome formation and degradation, (ii) the fusing of the autophagosome and mycobacterial phagosome followed by maturation or (iii) the activation of the endosomal system during autophagy induction ³¹¹. In addition, IFN- γ , an important cytokine in the fight against invading pathogens, is also able to induce autophagy in macrophages ³¹³. It is therefore conceivable that ANXA11, which has known vesicle trafficking properties ²⁷⁷, could also have a role in this pathway.

Table 18: Comparison of the allele frequencies in the German, South African Coloured and Yoruba control populations.

SNP	Allele frequencies in Germans ²⁷⁷			Allele frequencies in South African Coloured			Allele frequencies in Yoruba [§]		
	Allele	Allele Frequency	Heterozygote Frequency	Allele	Allele Frequency	Heterozygote Frequency	Allele	Allele Frequency	Heterozygote Frequency
rs2789679	T	0.44	0.48	T	0.23	0.35	T	0.17	0.33
	A	0.56		A	0.77		A	0.83	
rs7071579	A	0.57	0.50	A	0.48	0.45	A	0.47	0.5
	G	0.43		G	0.52		G	0.53	
rs1953600	T	0.45	0.51	T	0.26	0.38	T	0.17	0.33
	C	0.55		C	0.74		C	0.83	
rs2573346	A	0.45	0.51	A	0.23	0.35	A	0.12	0.24
	G	0.55		G	0.77		G	0.88	
rs2784773	T	0.36	0.48	T	0.45	0.51	T	0.49	0.52
	C	0.64		C	0.55		C	0.51	
rs1049550	A	0.41	0.49	A	0.26	0.37	A	0.17	0.33
	G	0.59		G	0.74		G	0.83	

[§]Allele frequencies for the Yoruba population were obtained from HapMap.

Chapter 6: CADM gene family

Knowing is not enough; we must apply. Willing is not enough; we must do.

Johann Wolfgang von Goethe

6.1 Cell Adhesion Molecules

6.1.1 Cell Adhesion Molecule Family

The immunoglobulin (Ig) superfamily of proteins is a member of the cell adhesion molecules (CAMs), and is the largest family with over 100 members already identified in vertebrates³¹⁶. These IgCAMs differ from other CAM families in that they are Ca²⁺-independent³¹⁶⁻³¹⁸. The IgCAMs consist of four main classes of molecules, including the NCAMs, L1 family CAMs and the nectins³¹⁶. All nectins are comprised of an extracellular region with three Ig-like loops in addition to a transmembrane- and a cytoplasmic domain³¹⁹. These nectin molecules are able to form homo-dimers in both the *cis* and *trans* formations which facilitates cell-cell adhesion. They are also able to form hetero-trans-dimers with other molecules. Recently, five nectin-like (NECL) molecules have been identified based on their similar structure to nectins. These include NECL-1; -2; -3; -4 and -5 (Table 19). In this thesis, these genes will be referred to as CADM1, CADM2 and CADM3. These molecules play various roles in cell-cell interactions, synapse formations, target-cell recognition, neuronal cell migration and the formation of complex glial networks^{316, 318, 320}.

Table 19: Various names used in the literature for the CADM genes.

CADM1	CADM2	CADM3
NECL-2	NECL-3	NECL-1
TSLC1	TSLL2	TSLL1
SynCAM1	SynCAM2	SynCAM3
SgIGSF		
IGSF4	IGSF4D	IGSF4B
sTSLC-1		
RA175		

6.1.2 CADM1

CADM1 is a type-1 transmembrane glycoprotein located on chromosome 11q23.2 and is highly conserved³²⁰⁻³²³. This membrane glycoprotein contains three Ig-like C2-type domains, a transmembrane domain and a short cytoplasmic domain (Figure 19)^{316, 321, 322}. This cytoplasmic domain contains two interaction motifs, namely the 4.1/ezrin/radixin/moesin (FERM) domain and the PDZ domain^{319, 322, 323} which acts as the docking sites for various transmembrane proteins³²³ conferring upon CADM1 its ability to play a role in cell-cell interactions³¹⁶. CADM1 is also present on the surface of dendritic cells and acts as a ligand for the class-I-restricted T-cell-associated molecule (CRTAM). Their interaction results in the release of cytokines³²³. Expression of *CADM1* has also been shown to be ubiquitous³²².

6.1.3 Other CADM Genes: CADM2, CADM3 and NCAM2

Not much is known about *CADM2* and *CADM3* genes. However, Fukhara *et al.* (2001) and Fukami *et al.* (2003) have shown great homology between the gene structure of *CADM1*, *CADM2* and *CADM3*, which allowed them to classify these three genes into a unique subfamily. Due to these similarities it is also believed that *CADM2* and *CADM3* play a role in cell adhesion. Interestingly however, the expression patterns of these genes differ markedly, with *CADM2* expressed restrictively in the prostate, brain and kidney; while *CADM3* is

exclusively expressed in the brain^{318, 322}. *Neural Cell Adhesion Molecule 2 (NCAM2)* gene structure was also significantly homologous to that of *CADM1*, and as mentioned above is a member of the Ig superfamily of proteins³²².

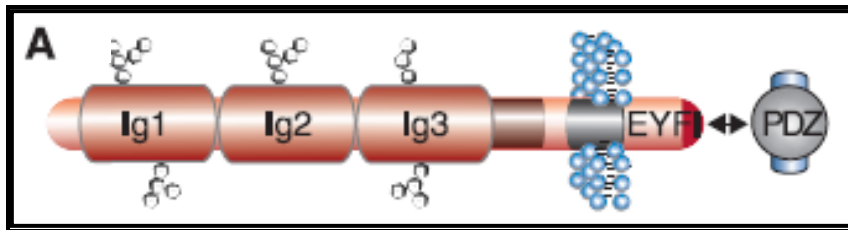


Figure 19: Structure of CADM1 protein, showing (left to right) the 3 Ig domains; the transmembrane domain and the short cytoplasmic domain³²⁰. The PDZ (post synaptic density protein (PSD95), *Drosophila* disc large tumour suppressor (DlgA), and zonula occludens-1 protein (zo-1)) domain is a common structural domain found in signalling proteins.

6.1.4 Disease Mechanisms – CADM family

6.1.4.1 Cancers

CADM1 has been shown to play a role in various cancers, including non-small cell lung cancer, hepatocellular carcinoma, pancreatic cancer, prostate cancer, breast cancer and gastric cancers^{317, 321, 322}. Various studies have shown that truncation of the cytoplasmic domain of CADM1 results in tumour formation and that restoration of CADM1 protein levels to normal results in tumour suppression^{324, 325}, and that the loss of one allele and suppression of the other allele by promoter methylation or mutation is frequently observed in the various cancers mentioned above^{321, 326, 327}. These studies have thus highlighted the tumour suppressor ability of CADM1. The study conducted by Fukhara *et al.* (2001) also showed that lower expression of *CADM2* and *CADM3* in various human glioma and prostate cancer cell lines is indicative of the important role of these genes in tumour suppression. However, the mechanism whereby this unique gene family is able to suppress tumour formation still eludes us.

6.1.4.2 TB

In a recent GWAS for TB in West Africans, there was an indication that *CADM1* may alter susceptibility to TB (unpublished). Polymorphisms associated with other CADM members and the highly homologous *NCAM2* gene were also found to play a role in TB susceptibility. The authors postulated that CADM1 could play an important role in antigen cross-presentation between dendritic cells (DC) and natural killer (NK) cells. CADM1 has been found to be present on the surface of a subset of DCs and through its interaction with CRTAM on the surface of NK cells; the cytokine IL-22 is released³²³. IL-22 is known to induce the epithelial innate immune response in the respiratory system, resulting in the release of anti-microbials including beta-defensins, which have been shown to have anti-mycobacterial activity. Therefore by altering NK and CD8 T cell activity, CADM1 could play a role in TB susceptibility by enhancing the mucosal innate immunity.

The Wellcome Trust Case Control Consortium (WTCCC) and The African Tuberculosis Genetics Group conducted the first two independent GWAS for TB, and subsequently the data was pooled, analysed and recently published⁵⁵. One of the draw-cards for GWAS is

that they are not hypothesis based, and therefore allow for the identification of novel genes and pathways. In the early analysis of the WTCCC TB GWAS, *CADM1* polymorphisms were found to alter susceptibility in West Africans (unpublished). Polymorphisms in other *CADM* genes, *CADM2* and *CADM3*, were also found to be associated with susceptibility to TB (Table 20). We were therefore interested in identifying whether or not these genes, previously not considered as susceptibility factors, could alter susceptibility to TB in the SAC population.

Table 20: *CADM* polymorphisms investigated as TB susceptibility variants in the SAC population.

Gene	Polymorphism	Location
CADM1	rs1460911	5' UTR/upstream
	rs1563899	5' UTR/upstream
	rs1563900	5' UTR/upstream
	rs2446890	5' UTR/upstream
	rs2515327	5' UTR/upstream
CADM2	rs2324979	3' UTR/downstream
CADM3	rs12057331	5' UTR/upstream
	rs12068892	5' UTR/upstream
	rs16841729	5' UTR/upstream
NCAM2	rs8184921	intron 1

6.2 Results

6.2.1 Genotype Analysis

6.2.1.1 *CADM1*

All of the polymorphisms investigated by TaqMan® genotyping had a call rate $\geq 95\%$ and were in HWE. $P < 0.05$ was considered statistically significant. However, none of the polymorphisms that were investigated were found to be significantly associated with susceptibility to TB (Table 21).

Table 21: Statistical analysis of *CADM1* polymorphisms.

SNP	Allele		Genotype frequencies								P_{gen}^{\dagger}	P_{allele}^*	OR [95%CI] [§]
			Cases				Controls						
	1 [‡]	2	11	12	22	HWE [#]	11	12	22	HWE			
rs1460911	T	C	0.22	0.47	0.31	0.26	0.20	0.48	0.32	0.69	0.72	0.52	0.94 [0.77-1.14]
rs1563899	A	G	0.16	0.46	0.38	0.11	0.15	0.42	0.43	0.053	0.44	0.41	1.09 [0.89-1.33]
rs1563900	C	T	0.16	0.46	0.38	0.52	0.16	0.42	0.42	0.054	0.53	0.54	1.07 [0.87-1.31]
rs2446890	G	A	0.14	0.46	0.40	0.74	0.15	0.41	0.44	0.058	0.37	0.38	0.91 [0.74-1.12]
rs2515327	C	A	0.11	0.45	0.44	0.78	0.12	0.39	0.49	0.067	0.15	0.52	0.93 [0.75-1.15]

[‡]Minor allele in controls.

[†] P value from a genotype-based X^2 -test, significant if $P < 0.05$.

^{*} P value from an allele-based Fisher's exact test, significant if $P < 0.05$.

[#]Hardy-Weinberg Equilibrium.

[§]Odds Ratio and [95% Confidence Interval]

6.2.1.2 *CADM2*

The *CADM2* polymorphism investigated by TaqMan® genotyping had a call rate of 96.4% and was in HWE. $P < 0.05$ was considered statistically significant. The polymorphism was however not associated with susceptibility to TB (Table 22).

Table 22: Statistical analysis of *CADM2* polymorphisms.

SNP	Allele		Genotype frequencies								P_{gen}^{\dagger}	P_{allele}^*	OR [95%CI] [§]
			Cases				Controls						
	1 [‡]	2	11	12	22	HWE [#]	11	12	22	HWE			
rs2324979	C	G	0.02	0.24	0.74	0.76	0.01	0.25	0.74	0.17	0.69	1	1.01 [0.76-1.34]

[‡]Minor allele in controls.

[†] P value from a genotype-based X^2 -test, significant if $P < 0.05$.

^{*} P value from an allele-based Fisher's exact test, significant if $P < 0.05$.

[#]Hardy-Weinberg Equilibrium.

[§]Odds Ratio and [95% Confidence Interval]

6.2.1.3 *CADM3*

Three *CADM3* polymorphisms were investigated using TaqMan® genotyping, with call rates $\geq 92\%$ and all were in HWE. $P < 0.05$ was considered statistically significant. Although none of the genotypes were found to be significantly associated with TB susceptibility, we did find a weak association with the G-allele of rs12057331 (OR 0.70, 95% CI 0.49-0.98) and susceptibility to TB. The two remaining polymorphisms were not found to be significantly associated with TB susceptibility (Table 23).

Table 23: Statistical analysis of *CADM3* polymorphisms.

SNP	Allele		Genotype frequencies								P_{gen}^{\dagger}	P_{allele}^*	OR [95%CI] [§]
			Cases				Controls						
	1 [‡]	2	11	12	22	HWE [#]	11	12	22	HWE			
rs12057331	G	A	0.01	0.20	0.79	0.19	0.01	0.14	0.85	0.83	0.08	0.04	0.70 [0.49-0.98]
rs12068892	T	C	0.01	0.20	0.79	0.18	0.01	0.14	0.85	0.83	0.12	0.06	0.72 [0.50-1.02]
rs16841729	A	G	0.01	0.20	0.79	0.22	0.01	0.14	0.85	0.87	0.11	0.05	1.42 [1.00-2.01]

[‡]Minor allele in controls.

[†] P value from a genotype-based X^2 -test, significant if $P < 0.05$.

^{*} P value from an allele-based Fisher's exact test, significant if $P < 0.05$.

[#]Hardy-Weinberg Equilibrium.

[§]Odds Ratio and [95% Confidence Interval]

The rs12057331 polymorphism of *CADM3* was further investigated in susceptibility to TBM using TaqMan® genotyping and had a call rate of 90% and was in HWE. $P < 0.05$ was considered statistically significant. The polymorphism however was not associated with susceptibility to TBM (Table 24).

Table 24: Statistical analysis of the *CADM3* polymorphism, rs12057331, and susceptibility to TBM.

SNP	Allele		Genotype frequencies								P _{gen} [†]	P _{allele} [*]	OR [95%CI] [§]
			Cases				Controls						
	1 [‡]	2	11	12	22	HWE [#]	11	12	22	HWE			
rs12057331	G	A	0.03	0.23	0.74	0.56	0.03	0.18	0.79	0.30	0.82	0.27	0.64 [0.31-1.32]

[‡]Minor allele in controls.

[†] P value from a genotype-based X²-test, significant if P<0.05.

^{*} P value from an allele-based Fisher's exact test, significant if P<0.05.

[#]Hardy-Weinberg Equilibrium.

[§]Odds Ratio and [95% Confidence Interval]

6.2.1.4 *NCAM2*

The *NCAM2* polymorphism investigated by TaqMan[®] genotyping had a call rate of 98.5% and was in HWE. P < 0.05 was considered statistically significant. The polymorphism however was not associated with susceptibility to TB (Table 25).

Table 25: Statistical analysis of *NCAM2* polymorphisms.

SNP	Allele		Genotype frequencies								P _{gen} [†]	P _{allele} [*]	OR [95%CI] [§]
			Cases				Controls						
	1 [‡]	2	11	12	22	HWE [#]	11	12	22	HWE			
rs8184921	T	C	0.13	0.44	0.43	0.52	0.13	0.44	0.43	0.71	1	0.84	0.97 [0.80-1.20]

[‡]Minor allele in controls.

[†] P value from a genotype-based X²-test, significant if P<0.05.

^{*} P value from an allele-based Fisher's exact test, significant if P<0.05.

[#]Hardy-Weinberg Equilibrium.

[§]Odds Ratio and [95% Confidence Interval]

6.2.2 Linkage Disequilibrium and Haplotype Analysis

6.2.2.1 *CADM1*

The LD structure of the five *CADM1* polymorphisms were all shown to be in strong LD (Figure 20), as identified by Haploview. However, of the five proposed haplotypes identified by the Cocaphase program, none were found to be significantly associated with susceptibility to TB in the SAC population (Table 26).

6.2.2.2 *CADM3*

The three *CADM3* polymorphisms were found to be in complete LD (Figure 21) with each other. However, Cocaphase could not predict any haplotypes based on the data generated by typing of the *CADM3* polymorphisms.

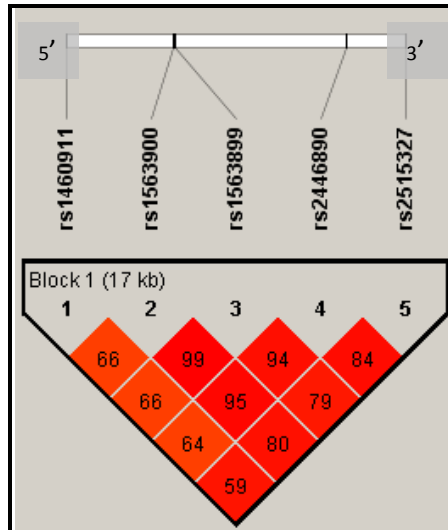


Figure 20: Plot of LD between the *CADM1* markers in the SAC population, generated by Haploview v4.1. The 5' and 3' ends of the genes are indicated and r^2 values (%) are indicated on the squares. The colours of the squares represent D' values, with red being $D' = 1$ and blue $D' = 0$.

Table 26: Haplotype analysis for *CADM1* polymorphisms.

Haplotypes*					Frequency		P value	OR
rs1460911	rs1563899	rs1563900	rs2446890	rs2515327	Cases	Controls		
C	G	T	A	A	0.53	0.55	0.44	0.56
T	A	C	A	A	0.09	0.01	0.58	0.44
T	A	C	G	A	0.031	0.027	0.61	0.66
T	A	C	G	C	0.33	0.32	0.7	0.59
T	G	T	A	A	0.08	0.08	0.95	0.59
Global significance#							0.171	

* Estimated haplotypes, as produced by Cocophase.

Permutation test P value calculated from 10 000 permutations in Cocophase, to correct for multiple testing while taking into account the correlation between markers and haplotypes.

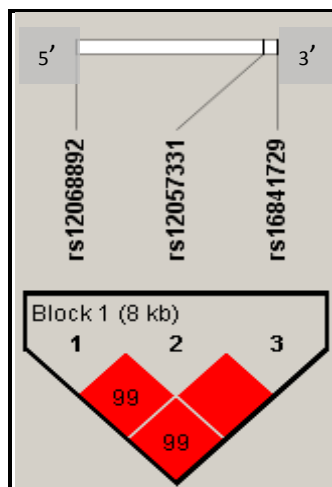


Figure 21: Plot of LD between the *CADM3* markers in the SAC population, generated by Haploview v4.1. The 5' and 3' ends of the genes are indicated and r^2 values (%) are indicated on the squares. The colours of the squares represent D' values, with red being $D' = 1$ and blue $D' = 0$.

6.3 Discussion

In the first TB GWAS conducted by The Wellcome Case Control Consortium polymorphisms associated with *CADM1* were found to alter susceptibility to TB in the Gambian population (unpublished). Polymorphisms associated with *CADM2*, *CADM3* and the structurally similar *NCAM2* genes were also found to be associated with susceptibility to TB during the “first stage” of genotyping in the GWAS. We were therefore interested in investigating whether or not these polymorphisms would also alter susceptibility to TB in the SAC population, in a TB endemic region.

In 2005, Boles *et al.* demonstrated the importance of *CADM1* in immuno-surveillance through its interaction with CRTAM³¹⁷. CRTAM had been found to be expressed on the surface of activated NK cells and T cells. This CRTAM-*CADM1* interaction was then subsequently shown to promote cytotoxicity of NK cells and CD8⁺ T cell secretion of IFN- γ in vitro. Therefore, it has been suggested that *CADM1* could act as a molecular target and play a role in distinguishing tumour cells from normal cells. In the TB GWAS, the authors postulated that this CRTAM-*CADM1* interaction could also alter susceptibility to TB. They argued that the binding of CRTAM to *CADM1* on antigen presenting cells results in the release of IL-22³²³. This cytokine is then able to induce the epithelial innate immunity by facilitating the secretion of anti-microbials (beta-defensins), some of which have defined anti-mycobacterial activity³²⁸⁻³³¹. Therefore, by enhancing the mucosal innate immunity *CADM1* could potentially alter susceptibility to TB.

In the TB GWAS that was conducted, five polymorphisms were found to alter susceptibility to TB in the Gambian population ($P = 1.38 \times 10^{-5} - 2.03 \times 10^{-6}$). However, during the replication phase of the GWAS, which was conducted in three populations; Guinea-Conakry, Guinea-Bassau and Malawi, only three SNPs identified during the initial genotyping phase were found to be significantly associated with TB susceptibility, and this was only true in the Guinea-Conakry population. In this study, we aimed to validate the findings of the initial genotyping phase in the SAC population. Here we did not find any of the polymorphisms identified in the TB GWAS to be significantly associated with susceptibility to TB. We also did not observe any association between the *CADM2* and *NCAM2* polymorphisms and susceptibility to TB.

On the other hand, for the *CADM3* polymorphisms, we did observe a weak association with the G-allele of rs12057331 polymorphism and susceptibility to TB. Haplotype analysis using Cocophase did not identify any haplotypes based on the *CADM3* genotyping data. However, we do not believe that this association is true but rather an artefact. Both Fukhara *et al.* (2001) and Fukhami *et al.* (2003) have shown that *CADM3* is expressed exclusively in brain and neurogenic cells; however, because an association was identified in the West African TB GWAS, we were interested in the *CADM* family. Since our TB cases were predominantly pulmonary TB individuals, the association identified becomes highly unlikely. We therefore investigated the rs12057331 polymorphism in a smaller cohort consisting of TBM cases. However, still no association was observed. After further investigation of polymorphism locations and the gene structure surrounding these SNPs, interesting observations were noted. Firstly, these polymorphisms occur in the 5' flanking region of the *CADM3* gene and secondly, they are approximately 50 kilo bases (kb) - 60 kb from either *CADM3* gene or the

absent in melanoma 2 (*AIM2*) gene. Also, the Duffy blood group, chemokine receptor (*DARC*) gene has been shown to overlap with the 3' region of *CADM3*. *AIM2* is a member of the IFI202X/IFI16 family and is known to play a role in tumourigenic reversion and may control cell proliferation. *AIM2* expression has also been shown to be induced by IFN- γ ³³². The *DARC* protein on the other hand is a glycosylated membrane protein and is a non-specific receptor for several chemokines. The encoded protein is also the receptor for the human malarial parasites, *Plasmodium vivax* and *Plasmodium knowlesi*³³³. *DARC* polymorphisms have also been shown to alter susceptibility to HIV/AIDS³³⁴. It is therefore likely that the association identified could hold true, if the polymorphisms in fact regulate either *AIM2* or *DARC*. Further research is therefore required.

The recently published GWAS that was conducted in two West African populations by the WTCCC and the African TB Genetics Consortium, identified a polymorphism, rs4331426 ($P = 6.8 \times 10^{-9}$, OR[95%] = 1.19[1.13 – 1.27]), in a gene poor region on chromosome 18q11.2 to be associated with susceptibility to TB⁵⁵. It is therefore evident that the *CADM1* locus that was identified in the initial WTCCC TB GWAS (unpublished) was no longer significantly associated with susceptibility to TB when the data of the African TB Genetics Consortium GWAS was included, indicating that it was probably an artifactual finding. It is therefore not surprising that our case-control study did not find association with this locus and the genes within it.

There are currently two schools of thought regarding the use and applicability of GWAS in elucidating novel and the true underlying disease genetic risk factors. On the one hand, due to its ability to interrogate the entire genome and its hypothesis-free basis, some researchers believe that the use of GWAS will lead to the discovery of the genetic factors that predispose individuals to disease, including complex diseases. Others however, believe that due to GWAS being driven by the CDCV hypothesis, they will not be able to successfully identify the genetic factors that predispose individuals to infectious diseases, as susceptibility to these diseases may be due to the presence of several rare variants. Based on the current literature, there is support for both schools of thought. GWAS have been successfully employed in the identification of susceptibility variants for various complex diseases, including Crohn's disease, diabetes and prostate cancer and even in some infectious diseases (HIV, Kawasaki disease, chronic hepatitis B and leprosy)^{103, 104, 107, 108, 335}. However, the current TB GWAS may lend support to those who do not believe that the CDCV hypothesis is appropriate. Another limiting factor for GWAS in African populations is the unavailability of microarray chips that contain SNP information relevant to these populations, as current microarray chips are based on European populations. It is known that African populations such as the West African populations used for the TB GWAS, have shorter stretches of LD, perhaps making them less likely to yield results in this type of enquiry³³⁶.

Therefore, until the current limitations of GWAS are addressed, GWAS may not be the best scientific tool to elucidate the genetic factors that alter susceptibility to TB and other methods such as admixture mapping, copy number variation and searching for rare variants may be better suited for the task at hand.

Chapter 7: Conclusions

Most great people have attained their greatest success just one step beyond their greatest failure.

Napoleon Hill

The work presented in this thesis provides further support for the hypothesis that host genetics plays an important role in TB disease outcome, and adds to the growing amount of evidence for this conclusion. In this study we investigated the role of five genes previously not considered to play a role in TB susceptibility.

Genome-wide association studies offer researchers a means to interrogate the entire genome for the presence of disease susceptibility factors, and are able to identify novel variants and pathways without being limited by our understanding of the disease pathology. However, their use in non-European populations, such as African populations who have a high disease burden, is still limited. The main reason for this is that the current SNP-chips used in GWAS are only able to cover approximately 46% of all SNPs currently available for the Yoruba population in the HapMap database.

ANXA11, a gene found in a GWAS on sarcoidosis, a granulomatous disease related to TB, also appeared to be implicated in TB, illustrating the possibility of using diseases with common pathways to aid in elucidating the causes of both. We identified the 3' UTR polymorphism in *ANXA11* as potentially altering the risk of developing TB. Previous studies on the role of *NRAMP1* have also identified polymorphisms in its 3' UTR that altered susceptibility to TB in West African and Asian populations, amongst others. Given the importance of the 3' UTR in post-transcriptional regulation due to the presence of a number of regulatory sequences (the polyadenylation signal and binding sites for various proteins and mRNA); it is not surprising that polymorphisms in this region are being implicated in disease susceptibility.

Whether or not apoptosis is beneficial to the host or the invading pathogen, the cellular process seems to be central to the outcome of disease. Another cellular process, autophagy, has also been shown to be important in the progress of disease, especially when the invading pathogen is able to suppress the apoptosis pathway. Future research should focus on the various molecules involved in these processes and how they affect susceptibility to TB.

In summary, an association between a 3' UTR *ANXA11* polymorphism that could alter susceptibility to TB was shown in this study, the biological mechanism of which still eludes us. It is possible that this polymorphism is not the actual causal variant but is in LD with it. It would therefore be important to genotype other polymorphisms in the 3' region of *ANXA11*, including the fifteen novel SNPs identified in this study. Investigating variation in the promoter region of this gene could also yield interesting results. None of the polymorphisms investigated in the *CADM* genes were found to be associated with susceptibility to TB.

In conclusion, a better understanding of the role that hosts genetics plays in disease susceptibility can only aid us in our fight of eradicating TB disease from society. The concept of genetically-based personalized medicine is fast becoming a reality. However, translation of the findings in this field to clinically relevant information is still in the future. Understanding how host genetics influences disease outcomes could allow us to identify those individuals who are at greater risk of developing disease. It could also allow us to design new vaccines and treatment regimens. For decades, TB has been treated by the use of various antibiotics, with no true success in eradicating the disease. Maybe a new tactic is

required, one that involves boosting the immune response in those individuals who seem to be unable to overcome the disease, leading to better results.

References

1. Daniel TM (2006) The history of tuberculosis. *Respir Med* 100 (11):1862-1870
2. Mathema B, Kurepina NE, Bifani PJ, and Kreiswirth BN (2006) Molecular epidemiology of tuberculosis: current insights. *Clin Microbiol Rev* 19 (4):658-685
3. Leão SC and Portaels F (2007) History. In: Palomino JC, Leão SC, Ritacco V (eds) *Tuberculosis 2007: From Basic Science to Patient Care.*, pp 25-49
4. Pease AS (1940) Some remarks on the diagnosis and treatment of tuberculosis in antiquity. *Isis* 31:380-393
5. Bates JH and Stead WW (1993) The history of tuberculosis as a global epidemic. *Med Clin North Am* 77 (6):1205-1217
6. Daniel TM (1997) *Captain of Death: the story of Tuberculosis*. University of Rochester Press. New York,
7. Canetti G, Rist N, and Grosset J (1964) Primary drug resistance in tuberculosis. *Am Rev Respir Dis* 90:792-799
8. World Health Organization. WHO report 2009: Global Tuberculosis Control 2009 - epidemiology, strategy, financing. WHO/HTM/TB/2009.411. 2009. Geneva.
9. Health Statistics. 2008.
10. U.S.Global Health Policy. Adult HIV/AIDS Prevalence Rate (Aged 15-49) 2007-2008. <http://www.globalhealthfacts.org/topic.jsp?i=3> . 2009.
11. World Health Organization. WHO Report 2008: Global tuberculosis control - surveillance, planning, financing. WHO/HTM/TB/2008.393. 2008. Geneva.
12. Frieden TR, Sterling TR, Munsiff SS, Watt CJ, and Dye C (2003) Tuberculosis. *Lancet* 362 (9387):887-899
13. Greenwood D, Slack RCB, and Peutherer JF (1997) *Medical Microbiology - A guide to microbial infections: pathogenesis, immunity, laboratory diagnosis and control*. Churchill Livingstone, London
14. Knechel NA (2009) Tuberculosis: pathophysiology, clinical features, and diagnosis. *Crit Care Nurse* 29 (2):34-43
15. Kaufmann SH and McMichael AJ (2005) Annulling a dangerous liaison: vaccination strategies against AIDS and tuberculosis. *Nat Med* 11 (4 Suppl):S33-S44
16. Pieters J (2001) Entry and survival of pathogenic mycobacteria in macrophages. *Microbes Infect* 3 (3):249-255
17. Smith I (2003) Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence. *Clin Microbiol Rev* 16 (3):463-496

18. Frieden TR, Sterling TR, Munsiff SS, Watt CJ, and Dye C (2003) Tuberculosis. *Lancet* 362 (9387):887-899
19. Kaufmann SH (2001) How can immunology contribute to the control of tuberculosis? *Nat Rev Immunol* 1 (1):20-30
20. Centre for Disease Control. TB Treatment. 2010.
21. Centers for Disease Control and Prevention (CDC) (2006) Emergence of *Mycobacterium tuberculosis* with extensive resistance to second-line drugs--worldwide, 2000-2004. *MMWR Morb Mortal Wkly Rep* 55 (11):301-305
22. Victor TC, Warren R, Beyers N, and van Helden PD (1997) Transmission of multidrug-resistant strains of *Mycobacterium tuberculosis* in a high incidence community. *Eur J Clin Microbiol Infect Dis* 16 (7):548-549
23. Victor TC, Streicher EM, Kewley C, Jordaan AM, van der Spuy GD, Bosman M, Louw H, Murray M, Young D, van Helden PD, and Warren RM (2007) Spread of an emerging *Mycobacterium tuberculosis* drug-resistant strain in the western Cape of South Africa. *Int J Tuberc Lung Dis* 11 (2):195-201
24. Johnson MD and Decker CF (2006) Tuberculosis and HIV Infection. *Dis Mon* 52 (11-12):420-427
25. Whalen C, Horsburgh CR, Hom D, Lahart C, Simberkoff M, and Ellner J (1995) Accelerated course of human immunodeficiency virus infection after tuberculosis. *Am J Respir Crit Care Med* 151 (1):129-135
26. Comstock GW (1978) Tuberculosis in twins: a re-analysis of the Proffit survey. *Am Rev Respir Dis* 117 (4):621-624
27. Kallmann FJ and Reisner D (1943) Twin studies on the significance of genetic factors in tuberculosis. *Am Rev Tuberc* 47:549-547
28. Simonds B (2004) Tuberculosis in twins. Pitman Medical Publishing Company,
29. Bellamy R (2003) Susceptibility to mycobacterial infections: the importance of host genetics. *Genes Immun* 4 (1):4-11
30. Hill AV (1999) Genetics and genomics of infectious disease susceptibility. *Br Med Bull* 55 (2):401-413
31. Hoal EG (2002) Human genetic susceptibility to tuberculosis and other mycobacterial diseases. *IUBMB Life* 53 (4-5):225-229
32. Möller M, de Wit E, and Hoal EG (2010) Past, present and future directions in human genetic susceptibility to tuberculosis. *FEMS Immunol Med Microbiol* 58 (1):3-26
33. Diehl K and Von Verschuer O (1936) Der Erbeinfluss bei der Tuberkulose. Gustav Fischer, London
34. Uehlinger E and Küinsch M (1938) Über Zwillingtuberkulose. *Beitrage zur Klinik der Tuberkulose* 92 (4):275-370

35. Harvald H and Hauge M (1965) Hereditary factors elucidated by twin studies. In: Neel JV, Shaw MW, Schull WL (eds) *Genetics and the epidemiology of chronic diseases*. S Department of Health, Education and Welfare, US Public Health Service, Division of Chronic Diseases, Washington, pp 61-76
36. van der Eijk EA, Van D, V, Vandenbroucke JP, and van Dissel JT (2007) Heredity versus environment in tuberculosis in twins: the 1950s United Kingdom Prophit Survey Simonds and Comstock revisited. *Am J Respir Crit Care Med* 176 (12):1281-1288
37. Alcaïs A and Abel L (2004) Application of genetic epidemiology to dissecting host susceptibility/resistance to infection illustrated with the study of common mycobacterial infections. In: Bellamy R (ed) *Susceptibility to infectious diseases - The importance of host genetics*. Cambridge University Press, Cambridge, pp 7-33
38. Motulsky AG (1960) Metabolic polymorphisms and the role of infectious diseases in human evolution. *Human Biology* 32:28-62
39. Serpa JA, Teeter LD, Musser JM, and Graviss EA (2009) Tuberculosis disparity between US-born blacks and whites, Houston, Texas, USA. *Emerg Infect Dis* 15 (6):899-904
40. Dubos R and Dubos J (1952) *The White Plague: Tuberculosis, Man and Society*. Little, Brown & Co., Boston
41. Farmer P, Robin S, Ramilus SL, and Kim JY (1991) Tuberculosis, poverty, and "compliance": lessons from rural Haiti. *Semin Respir Infect* 6 (4):254-260
42. Farmer P (1997) Social scientists and the new tuberculosis. *Soc Sci Med* 44 (3):347-358
43. Cegielski JP and McMurray DN (2004) The relationship between malnutrition and tuberculosis: evidence from studies in humans and experimental animals. *Int J Tuberc Lung Dis* 8 (3):286-298
44. Schaible UE and Kaufmann SH (2007) Malnutrition and infection: complex mechanisms and global impacts. *PLoS Med* 4 (5):e115
45. Finch PJ, Millard FJ, and Maxwell JD (1991) Risk of tuberculosis in immigrant Asians: culturally acquired immunodeficiency? *Thorax* 46 (1):1-5
46. Strachan DP, Powell KJ, Thaker A, Millard FJ, and Maxwell JD (1995) Vegetarian diet as a risk factor for tuberculosis in immigrant south London Asians. *Thorax* 50 (2):175-180
47. Chanarin I and Stephenson E (1988) Vegetarian diet and cobalamin deficiency: their association with tuberculosis. *J Clin Pathol* 41 (7):759-762
48. Finch PJ, Ang L, Eastwood JB, and Maxwell JD (1992) Clinical and histological spectrum of osteomalacia among Asians in south London. *Q J Med* 83 (302):439-448
49. Yang S, Smith C, Prah JM, Luo X, and Deluca HF (1993) Vitamin D deficiency suppresses cell-mediated immunity in vivo. *Arch Biochem Biophys* 303 (1):98-106
50. Rehm J and Parry CD. Alcohol consumption and infectious diseases in South Africa. *The Lancet* (in press)

51. Rehm J, Samokhvalov AV, Neuman MG, Room R, Parry C, Lonroth K, Patra J, Poznyak V, and Popova S (2009) The association between alcohol use, alcohol use disorders and tuberculosis (TB). A systematic review. *BMC Public Health* 9:450
52. Rehm J, Baliunas D, Borges GL, Graham K, Irving H, Kehoe T, Parry CD, Patra J, Popova S, Poznyak V, Roerecke M, Room R, Samokhvalov AV, and Taylor B (2010) The relation between different dimensions of alcohol consumption and burden of disease: an overview. *Addiction* 105 (5):817-843
53. Siddiqi K, Khan A, Ahmad M, and Rehman SU (2010) An intervention to stop smoking among patients suspected of TB - evaluation of an integrated approach. *BMC Public Health* 10 (1):160
54. Clementi M and Di Gianantonio E (2006) Genetic susceptibility to infectious diseases. *Reprod Toxicol* 21 (4):345-349
55. Thye T, Vannberg FO, Wong SH, Owusu-Dabo E, Osei I, Gyapong J, Sirugo G et al (2010) Genome-wide association analyses identifies a susceptibility locus for tuberculosis on chromosome 18q11.2. *Nat Genet* 42 (9):739-741
56. Richards EJ (2006) Inherited epigenetic variation--revisiting soft inheritance. *Nat Rev Genet* 7 (5):395-401
57. Abel L and Dessein AJ (1997) The impact of host genetics on susceptibility to human infectious diseases. *Curr Opin Immunol* 9 (4):509-516
58. Cervino AC, Lakiss S, Sow O, Bellamy R, Beyers N, Hoal-Van Helden E, van Helden P, McAdam KP, and Hill AV (2002) Fine mapping of a putative tuberculosis-susceptibility locus on chromosome 15q11-13 in African families. *Hum Mol Genet* 11 (14):1599-1603
59. Bellamy R, Beyers N, McAdam KP, Ruwende C, Gie R, Samaai P, Bester D, Meyer M, Corrah T, Collin M, Camidge DR, Wilkinson D, Hoal-Van Helden E, Whittle HC, Amos W, van Helden P, and Hill AV (2000) Genetic susceptibility to tuberculosis in Africans: a genome-wide scan. *Proc Natl Acad Sci U S A* 97 (14):8005-8009
60. Neyrolles O and Quintana-Murci L (2009) Sexual inequality in tuberculosis. *PLoS Med* 6 (12):e1000199
61. Davila S, Hibberd ML, Hari DR, Wong HE, Sahiratmadja E, Bonnard C, Alisjahbana B, Szeszko JS, Balabanova Y, Drobniowski F, van CR, Van D, V, Nejentsev S, Ottenhoff TH, and Seielstad M (2008) Genetic association and expression studies indicate a role of toll-like receptor 8 in pulmonary tuberculosis. *PLoS Genet* 4 (10):e1000218
62. Stein CM, Zalwango S, Malone LL, Won S, Mayanja-Kizza H, Mugerwa RD, Leontiev DV, Thompson CL, Cartier KC, Elston RC, Iyengar SK, Boom WH, and Whalen CC (2008) Genome scan of *M. tuberculosis* infection and disease in Ugandans. *PLoS ONE* 3 (12):e4094
63. Cobat A, Gallant CJ, Simkin L, Black GF, Stanley K, Hughes J, Doherty TM, Hanekom WA, Eley B, Jais JP, Boland-Auge A, van HP, Casanova JL, Abel L, Hoal EG, Schurr E, and Alcais A (2009) Two loci control tuberculin skin test reactivity in an area hyperendemic for tuberculosis. *J Exp Med* 206 (12):2583-2591

64. Möller M and Hoal EG (2010) Current findings, challenges and novel approaches in human genetic susceptibility to tuberculosis. *Tuberculosis (Edinb)*
65. Miller EN, Jamieson SE, Joberty C, Fakiola M, Hudson D, Peacock CS, Cordell HJ, Shaw MA, Lins-Lainson Z, Shaw JJ, Ramos F, Silveira F, and Blackwell JM (2004) Genome-wide scans for leprosy and tuberculosis susceptibility genes in Brazilians. *Genes Immun* 5 (1):63-67
66. Baghdadi JE, Orlova M, Alter A, Ranque B, Chentoufi M, Lazrak F, Archane MI, Casanova JL, Benslimane A, Schurr E, and Abel L (2006) An autosomal dominant major gene confers predisposition to pulmonary tuberculosis in adults. *J Exp Med* 203 (7):1679-1684
67. Cooke GS, Campbell SJ, Bennett S, Lienhardt C, McAdam KP, Sirugo G, Sow O, Gustafson P, Mwangulu F, van HP, Fine P, Hoal EG, and Hill AV (2008) Mapping of a novel susceptibility locus suggests a role for MC3R and CTSZ in human tuberculosis. *Am J Respir Crit Care Med* 178 (2):203-207
68. Mahasirimongkol S, Yanai H, Nishida N, Ridruechai C, Matsushita I, Ohashi J, Summanapan S, Yamada N, Moolphate S, Chuchotaworn C, Chairprasert A, Manosuthi W, Kantipong P, Kanitwittaya S, Sura T, Khusmith S, Tokunaga K, Sawanpanyalert P, and Keicho N (2009) Genome-wide SNP-based linkage analysis of tuberculosis in Thais. *Genes Immun* 10 (1):77-83
69. Lander ES and Schork NJ (1994) Genetic dissection of complex traits. *Science* 265 (5181):2037-2048
70. Burgner D, Jamieson SE, and Blackwell JM (2006) Genetic susceptibility to infectious diseases: big is beautiful, but will bigger be even better? *Lancet Infect Dis* 6 (10):653-663
71. Risch N and Merikangas K (1996) The future of genetic studies of complex human diseases. *Science* 273 (5281):1516-1517
72. Daly AK and Day CP (2001) Candidate gene case-control association studies: advantages and potential pitfalls. *Br J Clin Pharmacol* 52 (5):489-499
73. Wilkening S, Chen B, Bermejo JL, and Canzian F (2009) Is there still a need for candidate gene approaches in the era of genome-wide association studies? *Genomics* 93 (5):415-419
74. Gambaro G, Anglani F, and D'Angelo A (2000) Association studies of genetic polymorphisms and complex disease. *Lancet* 355 (9200):308-311
75. Ardlie KG, Kruglyak L, and Seielstad M (2002) Patterns of linkage disequilibrium in the human genome. *Nat Rev Genet* 3 (4):299-309
76. Newport MJ and Nejentsev S (2004) Genetics of susceptibility to tuberculosis in humans. *Monaldi Arch Chest Dis* 61 (2):102-111
77. Newport MJ (2009) Why hasn't human genetics told us more about tuberculosis? *Int J Tuberc Lung Dis* 13 (9):1049-1050
78. Spielman RS, McGinnis RE, and Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 52 (3):506-16

79. Lombard Z, Brune AE, Hoal EG, Babb C, van Helden PD, Epplen JT, and Bornman L (2006) HLA class II disease associations in southern Africa. *Tissue Antigens* 67 (2):97-110
80. Janeway C, Travers P, Walport M, and Shlomchik M (2001) *Immunobiology: the immune system in health and disease*. Fifth ed. Garland Science, New York
81. Khomenko AG, Litvinov VI, Chukanova VP, and Pospelov LE (1990) Tuberculosis in patients with various HLA phenotypes. *Tubercle* 71 (3):187-192
82. Meyer T, Baynes R, Bothwell T, Jenkins T, Jooste P, Dutoit E, Martell R, and Jacobs P (1990) Phenotypic-Expression of the Hla Linked Iron-Loading Gene in Males Over the Age of 40 Years - A Population Study Using Serial Serum Ferritin Estimations. *Journal of Internal Medicine* 227 (6):397-406
83. Ravikumar M, Dheenadhayalan V, Rajaram K, Lakshmi SS, Kumaran PP, Paramasivan CN, Balakrishnan K, and Pitchappan RM (1999) Associations of HLA-DRB1, DQB1 and DPB1 alleles with pulmonary tuberculosis in south India. *Tuber Lung Dis* 79 (5):309-317
84. Selvaraj P, Uma H, Reetha AM, Kurian SM, Xavier T, Prabhakar R, and Narayanan PR (1998) HLA antigen profile in pulmonary tuberculosis patients & their spouses. *Indian J Med Res* 107:155-158
85. Vejbaesya S, Chierakul N, Luangtrakool K, Srinak D, and Stephens HA (2002) Associations of HLA class II alleles with pulmonary tuberculosis in Thais. *Eur J Immunogenet* 29 (5):431-434
86. Malo D, Vogan K, Vidal S, Hu J, Cellier M, Schurr E, Fuks A, Bumstead N, Morgan K, and Gros P (1994) Haplotype mapping and sequence analysis of the mouse *Nramp* gene predict susceptibility to infection with intracellular parasites. *Genomics* 23 (1):51-61
87. North RJ, LaCourse R, Ryan L, and Gros P (1999) Consequence of *Nramp1* deletion to *Mycobacterium tuberculosis* infection in mice. *Infect Immun* 67 (11):5811-5814
88. Vidal SM, Malo D, Vogan K, Skamene E, and Gros P (1993) Natural resistance to infection with intracellular parasites: isolation of a candidate for *Bcg*. *Cell* 73 (3):469-485
89. Bellamy R, Ruwende C, Corrah T, McAdam KP, Whittle HC, and Hill AV (1998) Variations in the *NRAMP1* gene and susceptibility to tuberculosis in West Africans. *N Engl J Med* 338 (10):640-644
90. Hoal EG, Lewis L-A, Jamieson SE, Tanzer F, Rossouw M, Victor T, Hillerman R, Beyers N, Blackwell JM, and van Helden PD (2004) *SLC11A1 (NRAMP1)* but not *SLC11A2 (NRAMP2)* polymorphisms are associated with susceptibility to tuberculosis in a high-incidence community in South Africa. *Int J Tuberc Lung Dis* 8 (12):1464-1471
91. Li HT, Zhang TT, Zhou YQ, Huang QH, and Huang J (2006) *SLC11A1 (formerly NRAMP1)* gene polymorphisms and tuberculosis susceptibility: a meta-analysis. *Int J Tuberc Lung Dis* 10 (1):3-12
92. Holmskov U, Thiel S, and Jensenius JC (2003) Collections and ficolins: humoral lectins of the innate immune defense. *Annu Rev Immunol* 21:547-578

93. Turner MW (2003) The role of mannose-binding lectin in health and disease. *Mol Immunol* 40 (7):423-429
94. Summerfield JA, Sumiya M, Levin M, and Turner MW (1997) Association of mutations in mannose binding protein gene with childhood infection in consecutive hospital series. *BMJ* 314 (7089):1229-1232
95. Hoal-van Helden EG, Epstein J, Victor TC, Hon D, Lewis LA, Beyers N, Zurakowski D, Ezekowitz AB, and van Helden PD (1999) Mannose-binding protein B allele confers protection against tuberculous meningitis. *Pediatr Res* 45 (4 Pt 1):459-464
96. Cosar H, Ozkinay F, Onay H, Bayram N, Bakiler AR, Anil M, Can D, and Ozkinay C (2008) Low levels of mannose-binding lectin confers protection against tuberculosis in Turkish children. *Eur J Clin Microbiol Infect Dis* 27 (12):1165-1169
97. Soborg C, Madsen HO, Andersen AB, Lillebaek T, Kok-Jensen A, and Garred P (2003) Mannose-binding lectin polymorphisms in clinical tuberculosis. *J Infect Dis* 188 (5):777-782
98. The International HapMap Consortium (2003) The International HapMap Project. *Nature* 426 (6968):789-796
99. The International HapMap Consortium (2005) A haplotype map of the human genome. *Nature* 437 (7063):1299-1320
100. The International HapMap Consortium (2007) A second generation human haplotype map of over 3.1 million SNPs. *Nature* 449 (7164):851-861
101. Balding DJ (2006) A tutorial on statistical methods for population association studies. *Nat Rev Genet* 7 (10):781-791
102. Hirschhorn JN and Daly MJ (2005) Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 6 (2):95-108
103. Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, Zhang K et al (2007) A whole-genome association study of major determinants for host control of HIV-1. *Science* 317 (5840):944-947
104. Burgner D, Davila S, Breunis WB, Ng SB, Li Y, Bonnard C, Ling L, Wright VJ, Thalamuthu A, Odam M, Shimizu C, Burns JC, Levin M, Kuijpers TW, and Hibberd ML (2009) A genome-wide association study identifies novel and functionally related susceptibility Loci for Kawasaki disease. *PLoS Genet* 5 (1):e1000319
105. Kamatani Y, Wattanapokayakit S, Ochi H, Kawaguchi T, Takahashi A, Hosono N, Kubo M, Tsunoda T, Kamatani N, Kumada H, Puseenam A, Sura T, Daigo Y, Chayama K, Chantratita W, Nakamura Y, and Matsuda K (2009) A genome-wide association study identifies variants in the HLA-DP locus associated with chronic hepatitis B in Asians. *Nat Genet* 41 (5):591-595
106. Wong SH, Hill AV, and Vannberg FO (2010) Genomewide association study of leprosy. *N Engl J Med* 362 (15):1446-1447
107. Davila S, Wright VJ, Khor CC, Sim KS, Binder A, Breunis WB, Inwald D, Nadel S, Betts H, Carrol ED, de GR, Hermans PW, Hazelzet J, Emonts M, Lim CC, Kuijpers TW, Martinon-

- Torres F, Salas A, Zenz W, Levin M, and Hibberd ML (2010) Genome-wide association study identifies variants in the CFH region associated with host susceptibility to meningococcal disease. *Nat Genet* 42 (9):772-776
108. Zhang FR, Huang W, Chen SM, Sun LD, Liu H, Li Y, Cui Y et al (2009) Genomewide association study of leprosy. *N Engl J Med* 361 (27):2609-2618
 109. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, and Altshuler D (2002) The structure of haplotype blocks in the human genome. *Science* 296 (5576):2225-9
 110. Jeffreys AJ, Kauppi L, and Neumann R (2001) Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. *Nat Genet* 29 (2):217-222
 111. Patil N, Berno AJ, Hinds DA, Barrett WA, Doshi JM, Hacker CR, Kautzer CR, Lee DH, Marjoribanks C, McDonough DP, Nguyen BT, Norris MC, Sheehan JB, Shen N, Stern D, Stokowski RP, Thomas DJ, Trulson MO, Vyas KR, Frazer KA, Fodor SP, and Cox DR (2001) Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science* 294 (5547):1719-1723
 112. Levenstien MA, Ott J, and Gordon D (2006) Are molecular haplotypes worth the time and expense? A cost-effective method for applying molecular haplotypes. *PLoS Genet* 2 (8):e127
 113. Abecasis GR, Ghosh D, and Nichols TE (2005) Linkage disequilibrium: ancient history drives the new genetics. *Hum Hered* 59 (2):118-124
 114. Weiss KM and Clark AG (2002) Linkage disequilibrium and the mapping of complex human traits. *Trends Genet* 18 (1):19-24
 115. Flynn JL (2006) Lessons from experimental *Mycobacterium tuberculosis* infections. *Microbes Infect* 8 (4):1179-1188
 116. Vidal S, Tremblay ML, Govoni G, Gauthier S, Sebastiani G, Malo D, Skamene E, Olivier M, Jothy S, and Gros P (1995) The *Ity/Lsh/Bcg* locus: natural resistance to infection with intracellular parasites is abrogated by disruption of the *Nramp1* gene. *J Exp Med* 182 (3):655-666
 117. Cellier M, Govoni G, Vidal S, Kwan T, Groulx N, Liu J, Sanchez F, Skamene E, Schurr E, and Gros P (1994) Human natural resistance-associated macrophage protein: cDNA cloning, chromosomal mapping, genomic organization, and tissue-specific expression. *J Exp Med* 180 (5):1741-1752
 118. Flynn JL, Goldstein MM, Chan J, Triebold KJ, Pfeffer K, Lowenstein CJ, Schreiber R, Mak TW, and Bloom BR (1995) Tumor necrosis factor- α is required in the protective immune response against *Mycobacterium tuberculosis* in mice. *Immunity* 2 (6):561-572
 119. Cooper AM, Dalton DK, Stewart TA, Griffin JP, Russell DG, and Orme IM (1993) Disseminated tuberculosis in interferon gamma gene-disrupted mice. *J Exp Med* 178 (6):2243-2247

120. Flynn JL, Chan J, Triebold KJ, Dalton DK, Stewart TA, and Bloom BR (1993) An essential role for interferon gamma in resistance to *Mycobacterium tuberculosis* infection. *J Exp Med* 178 (6):2249-2254
121. Kamijo R, Le J, Shapiro D, Havell EA, Huang S, Aguet M, Bosland M, and Vilcek J (1993) Mice that lack the interferon-gamma receptor have profoundly altered responses to infection with *Bacillus Calmette-Guerin* and subsequent challenge with lipopolysaccharide. *J Exp Med* 178 (4):1435-1440
122. Flynn JL, Goldstein MM, Triebold KJ, Koller B, and Bloom BR (1992) Major histocompatibility complex class I-restricted T cells are required for resistance to *Mycobacterium tuberculosis* infection. *Proc Natl Acad Sci U S A* 89 (24):12013-12017
123. Ladel CH, Daugelat S, and Kaufmann SH (1995) Immune response to *Mycobacterium bovis* bacille Calmette Guerin infection in major histocompatibility complex class I- and II-deficient knock-out mice: contribution of CD4 and CD8 T cells to acquired resistance. *Eur J Immunol* 25 (2):377-384
124. Ladel CH, Hess J, Daugelat S, Mombaerts P, Tonegawa S, and Kaufmann SH (1995) Contribution of alpha/beta and gamma/delta T lymphocytes to immunity against *Mycobacterium bovis* bacillus Calmette Guerin: studies with T cell receptor-deficient mutant mice. *Eur J Immunol* 25 (3):838-846
125. Kamijo R, Harada H, Matsuyama T, Bosland M, Gerecitano J, Shapiro D, Le J, Koh SI, Kimura T, Green SJ, and . (1994) Requirement for transcription factor IRF-1 in NO synthase induction in macrophages. *Science* 263 (5153):1612-1615
126. Ladel CH, Blum C, Dreher A, Reifenberg K, Kopf M, and Kaufmann SH (1997) Lethal tuberculosis in interleukin-6-deficient mutant mice. *Infect Immun* 65 (11):4843-4849
127. Cooper AM, Magram J, Ferrante J, and Orme IM (1997) Interleukin 12 (IL-12) is crucial to the development of protective immunity in mice intravenously infected with mycobacterium tuberculosis. *J Exp Med* 186 (1):39-45
128. Sanjeevi CB, Narayanan PR, Prabakar R, Charles N, Thomas BE, Balasubramaniam R, and Olerup O (1992) No association or linkage with HLA-DR or -DQ genes in south Indians with pulmonary tuberculosis. *Tuber Lung Dis* 73 (5):280-284
129. Teran-Escandon D, Teran-Ortiz L, Camarena-Olvera A, Gonzalez-Avila G, Vaca-Marin MA, Granados J, and Selman M (1999) Human leukocyte antigen-associated susceptibility to pulmonary tuberculosis: molecular analysis of class II alleles by DNA amplification and oligonucleotide hybridization in Mexican patients. *Chest* 115 (2):428-433
130. Amirzargar AA, Yalda A, Hajabolbaghi M, Khosravi F, Jabbari H, Rezaei N, Niknam MH, Ansari B, Moradi B, and Nikbin B (2004) The association of HLA-DRB, DQA1, DQB1 alleles and haplotype frequency in Iranian patients with pulmonary tuberculosis. *Int J Tuberc Lung Dis* 8 (8):1017-1021
131. Dubaniewicz A, Moszkowska G, Szczerkowska Z, and Hoppe A (2003) Analysis of DQB1 allele frequencies in pulmonary tuberculosis: preliminary report. *Thorax* 58 (10):890-891

132. Lombard Z, Dalton DL, Venter PA, Williams RC, and Bornman L (2006) Association of HLA-DR, -DQ, and vitamin D receptor alleles and haplotypes with tuberculosis in the Venda of South Africa. *Hum Immunol* 67 (8):643-654
133. Goldfeld AE, Delgado JC, Thim S, Bozon MV, Ugialoro AM, Turbay D, Cohen C, and Yunis EJ (1998) Association of an HLA-DQ allele with clinical tuberculosis. *JAMA* 279 (3):226-228
134. Delgado JC, Baena A, Thim S, and Goldfeld AE (2006) Aspartic acid homozygosity at codon 57 of HLA-DQ beta is associated with susceptibility to pulmonary tuberculosis in Cambodia. *J Immunol* 176 (2):1090-1097
135. Bothamley GH, Beck JS, Schreuder GM, D'Amato J, de Vries RR, Kardjito T, and Ivanyi J (1989) Association of tuberculosis and M. tuberculosis-specific antibody levels with HLA. *J Infect Dis* 159 (3):549-555
136. Singh SP, Mehra NK, Dingley HB, Pande JN, and Vaidya MC (1983) Human leukocyte antigen (HLA)-linked control of susceptibility to pulmonary tuberculosis and association with HLA-DR types. *J Infect Dis* 148 (4):676-681
137. Brahmajothi V, Pitchappan RM, Kakkanaiah VN, Sashidhar M, Rajaram K, Ramu S, Palanimurugan K, Paramasivan CN, and Prabhakar R (1991) Association of pulmonary tuberculosis and HLA in south India. *Tubercle* 72 (2):123-132
138. Rajalingam R, Singal DP, and Mehra NK (1997) Transporter associated with antigen-processing (TAP) genes and susceptibility to tuberculoid leprosy and pulmonary tuberculosis. *Tissue Antigens* 49 (2):168-172
139. Pospelov LE, Matrakshin AG, Chernousova LN, Tsoi KN, Afanasjev KI, Rubtsova GA, and Yeremeyev VV (1996) Association of various genetic markers with tuberculosis and other lung diseases in Tuvian children. *Tuber Lung Dis* 77 (1):77-80
140. Cox RA, Downs M, Neimes RE, Ognibene AJ, Yamashita TS, and Ellner JJ (1988) Immunogenetic analysis of human tuberculosis. *J Infect Dis* 158 (6):1302-1308
141. Chandanayingyong D, Maranetra N, and Bovornkitti S (1988) HLA antigen profiles in Thai tuberculosis patients. *Asian Pac J Allergy Immunol* 6 (2):77-80
142. Ruggiero G, Cosentini E, Zanzi D, Sanna V, Terrazzano G, Matarese G, Sanduzzi A, Perna F, and Zappacosta S (2004) Allelic distribution of human leukocyte antigen in historical and recently diagnosed tuberculosis patients in Southern Italy. *Immunology* 111 (3):318-322
143. Pospelova LE, Matrashkin AG, Larionova EE, Ereemeev VV, and Mes'ko EM (2005) [The association of tuberculosis with the specificities of the HLA gene DRB1 in different regions of Tuva]. *Probl Tuberk Bolezn Legk* (7):23-25
144. Kim HS, Park MH, Song EY, Park H, Kwon SY, Han SK, and Shim YS (2005) Association of HLA-DR and HLA-DQ Genes With Susceptibility to Pulmonary Tuberculosis in Koreans: Preliminary Evidence of Associations With Drug Resistance, Disease Severity, and Disease Recurrence. *Hum Immunol* 66 (10):1074-81
145. Wang J, Song C, and Wang S (2001) [Association of HLA-DRB1 genes with pulmonary tuberculosis]. *Zhonghua Jie He He Hu Xi Za Zhi* 24 (5):302-305

146. Dubaniewicz A, Lewko B, Moszkowska G, Zamorska B, and Stepinski J (2000) Molecular subtypes of the HLA-DR antigens in pulmonary tuberculosis. *Int J Infect Dis* 4 (3):129-133
147. Sriram U, Selvaraj P, Kurian SM, Reetha AM, and Narayanan PR (2001) HLA-DR2 subtypes & immune responses in pulmonary tuberculosis. *Indian J Med Res* 113:117-124
148. Blackwell JM (2001) Genetics and genomics in infectious disease susceptibility. *Trends Mol Med* 7 (11):521-526
149. Hawkins BR, Higgins DA, Chan SL, Lowrie DB, MITCHISON DA, and Girling DJ (1988) HLA typing in the Hong Kong Chest Service/British Medical Research Council study of factors associated with the breakdown to active tuberculosis of inactive pulmonary lesions. *Am Rev Respir Dis* 138 (6):1616-1621
150. Chu SF, Tam CM, Wong HS, Kam KM, Lau YL, and Chiang AK (2007) Association between RANTES functional polymorphisms and tuberculosis in Hong Kong Chinese. *Genes Immun* 8 (6):475-479
151. Möller M, Nebel A, Valentonyte R, van Helden PD, Schreiber S, and Hoal EG (2009) Investigation of chromosome 17 candidate genes in susceptibility to TB in a South African population. *Tuberculosis (Edinb)* 89 (2):189-194
152. Thye T, Nejentsev S, Intemann CD, Browne EN, Chinbuah MA, Gyapong J, Osei I, Owusu-Dabo E, Zeitels LR, Herb F, Horstmann RD, and Meyer CG (2009) MCP-1 promoter variant -362C associated with protection from pulmonary tuberculosis in Ghana, West Africa. *Hum Mol Genet* 18 (2):381-388
153. Flores-Villanueva PO, Ruiz-Morales JA, Song CH, Flores LM, Jo EK, Montano M, Barnes PF, Selman M, and Granados J (2005) A functional promoter polymorphism in monocyte chemoattractant protein-1 is associated with increased susceptibility to pulmonary tuberculosis. *J Exp Med* 202 (12):1649-1658
154. Gomez LM, Anaya JM, Sierra-Filardi E, Cadena J, Corbi A, and Martin J (2006) Analysis of DC-SIGN (CD209) functional variants in patients with tuberculosis. *Hum Immunol* 67 (10):808-811
155. Ben-Ali M, Barreiro LB, Chabbou A, Haltiti R, Braham E, Neyrolles O, Dellagi K, Gicquel B, Quintana-Murci L, and Barbouche MR (2007) Promoter and neck region length variation of DC-SIGN is not associated with susceptibility to tuberculosis in Tunisian patients. *Hum Immunol* 68 (11):908-912
156. Olesen R, Wejse C, Velez DR, Bisseye C, Sodemann M, Aaby P, Rabna P, Worwui A, Chapman H, Diatta M, Adegbola RA, Hill PC, Ostergaard L, Williams SM, and Sirugo G (2007) DC-SIGN (CD209), pentraxin 3 and vitamin D receptor gene variants associate with pulmonary tuberculosis risk in West Africans. *Genes Immun* 8 (6):456-467
157. Selvaraj P, Alagarasu K, Swaminathan S, Harishankar M, and Narendran G (2009) CD209 gene polymorphisms in South Indian HIV and HIV-TB patients. *Infect Genet Evol* 9 (2):256-262
158. Barreiro LB, Neyrolles O, Babb CL, Tailleux L, Quach H, McElreavey K, Helden PD, Hoal EG, Gicquel B, and Quintana-Murci L (2006) Promoter Variation in the DC-SIGN Encoding Gene CD209 Is Associated with Tuberculosis. *PLoS Medicine* 3 (2):e20

159. Vannberg FO, Chapman SJ, Khor CC, Tosh K, Floyd S, Jackson-Sillah D, Crampin A, Sichali L, Bah B, Gustafson P, Aaby P, McAdam KP, Bah-Sow O, Lienhardt C, Sirugo G, Fine P, and Hill AV (2008) CD209 genetic polymorphism and tuberculosis disease. *PLoS ONE* 3 (1):e1388
160. Tang NL, Fan HP, Chang KC, Ching JK, Kong KP, Yew WW, Kam KM, Leung CC, Tam CM, Blackwell J, and Chan CY (2009) Genetic association between a chemokine gene CXCL-10 (IP-10, interferon gamma inducible protein 10) and susceptibility to tuberculosis. *Clin Chim Acta* 406 (1-2):98-102
161. Martineau AR, Leandro AC, Anderson ST, Newton SM, Wilkinson KA, Nicol MP, Pienaar SM, Skolimowska KH, Rocha MA, Rolla VC, Levin M, Davidson RN, Bremner SA, Griffiths CJ, Eley BS, Bonecini-Almeida MG, and Wilkinson RJ (2009) Association between Gc genotype and susceptibility to TB is dependent on vitamin D status. *Eur Respir J* doi: 10.1183/09031936.00087009
162. Lio D, Marino V, Serauto A, Gioia V, Scola L, Crivello A, Forte GI, Colonna-Romano G, Candore G, and Caruso C (2002) Genotype frequencies of the +874T-->A single nucleotide polymorphism in the first intron of the interferon-gamma gene in a sample of Sicilian patients affected by tuberculosis. *Eur J Immunogenet* 29 (5):371-374
163. Etokebe GE, Bulat-Kardum L, Johansen MS, Knezevic J, Balen S, Matakovic-Mileusnic N, Matanic D, Flego V, Pavelic J, Beg-Zec Z, and Dembic Z (2006) Interferon-gamma gene (T874A and G2109A) polymorphisms are associated with microscopy-positive tuberculosis. *Scand J Immunol* 63 (2):136-141
164. Lopez-Maderuelo D, Arnalich F, Serantes R, Gonzalez A, Codoceo R, Madero R, Vazquez JJ, and Montiel C (2003) Interferon-gamma and interleukin-10 gene polymorphisms in pulmonary tuberculosis. *Am J Respir Crit Care Med* 167 (7):970-975
165. Cooke GS, Campbell SJ, Sillah J, Gustafson P, Bah B, Sirugo G, Bennett S, McAdam KP, Sow O, Lienhardt C, and Hill AV (2006) Polymorphism within the interferon-gamma/receptor complex is associated with pulmonary tuberculosis. *Am J Respir Crit Care Med* 174 (3):339-343
166. Ding S, Li L, and Zhu X (2008) Polymorphism of the interferon-gamma gene and risk of tuberculosis in a southeastern Chinese population. *Hum Immunol* 69 (2):129-133
167. Rossouw M, Nel HJ, Cooke GS, van Helden PD, and Hoal EG (2003) Association between tuberculosis and a polymorphic NFkappaB binding site in the interferon gamma gene. *Lancet* 361 (9372):1871-1872
168. Sallakci, Nilgun, Coskun, Mesut, Berber, Zafer, Gurkan, Fuat, Kocamaz, Halil, Uysal, Gulnar, Bhujju, Sabin, Yavuzer, Ugur, Singh, Mahavir, and Yegin, Olcay. Interferon- γ gene+874T-A polymorphism is associated with tuberculosis and gamma interferon response. *Tuberculosis* . 2007.
169. Vidyarani M, Selvaraj P, Prabhu AS, Jawahar MS, Adhilakshmi AR, and Narayanan PR (2006) Interferon gamma (IFN γ) & interleukin-4 (IL-4) gene variants & cytokine levels in pulmonary tuberculosis. *Indian J Med Res* 124 (4):403-410
170. Moran A, Ma X, Reich RA, and Graviss EA (2007) No association between the +874T/A single nucleotide polymorphism in the IFN-gamma gene and susceptibility to TB. *Int J Tuberc Lung Dis* 11 (1):113-115

171. Tso HW, Ip WK, Chong WP, Tam CM, Chiang AK, and Lau YL (2005) Association of interferon gamma and interleukin 10 genes with tuberculosis in Hong Kong Chinese. *Genes Immun* 6 (4):358-363
172. Henao MI, Montes C, Paris SC, and Garcia LF (2006) Cytokine gene polymorphisms in Colombian patients with different clinical presentations of tuberculosis. *Tuberculosis (Edinburgh, Scotland)* 86 (1):11-19
173. Pacheco AG, Cardoso CC, and Moraes MO (2008) IFNG +874T/A, IL10 -1082G/A and TNF -308G/A polymorphisms in association with tuberculosis susceptibility: a meta-analysis study. *Hum Genet* 123 (5):477-484
174. Akahoshi M, Ishihara M, Remus N, Uno K, Miyake K, Hirota T, Nakashima K, Matsuda A, Kanda M, Enomoto T, Ohno S, Nakashima H, Casanova JL, Hopkin JM, Tamari M, Mao XQ, and Shirakawa T (2004) Association between IFNA genotype and the risk of sarcoidosis. *Hum Genet* 114 (5):503-509
175. Sahiratmadja E, Baak-Pablo R, de Visser AW, Alisjahbana B, Adnan I, van Crevel R, Marzuki S, van Dissel JT, Ottenhoff TH, and Van D, V (2007) Association of polymorphisms in IL-12/IFN-gamma pathway genes with susceptibility to pulmonary tuberculosis in Indonesia. *Tuberculosis (Edinb)* 87 (4):303-311
176. Kusuhara K, Yamamoto K, Okada K, Mizuno Y, and Hara T (2007) Association of IL12RB1 polymorphisms with susceptibility to and severity of tuberculosis in Japanese: a gene-based association analysis of 21 candidate genes. *Int J Immunogenet* 34 (1):35-44
177. Thye T, Browne EN, Chinbuah MA, Gyapong J, Osei I, Owusu-Dabo E, Brattig NW, Niemann S, Rusch-Gerdes S, Horstmann RD, and Meyer CG (2009) IL10 haplotype associated with tuberculin skin test response but not with pulmonary TB. *PLoS ONE* 4 (5):e5420
178. Ansari A, Talat N, Jamil B, Hasan Z, Razzaki T, Dawood G, and Hussain R (2009) Cytokine gene polymorphisms across tuberculosis clinical spectrum in Pakistani patients. *PLoS ONE* 4 (3):e4778
179. Vollstedt S, Yuliwulandari R, Okamoto K, Lien LT, Keicho N, Rochani JT, Wikaningrum R, and Tokunaga K (2009) No evidence for association between the interferon regulatory factor 1 (IRF1) gene and clinical tuberculosis. *Tuberculosis (Edinb)* 89 (1):71-76
180. Soborg C, Andersen AB, Range N, Malenganisho W, Friis H, Magnussen P, Temu MM, Chungalucha J, Madsen HO, and Garred P (2007) Influence of candidate susceptibility genes on tuberculosis in a high endemic region. *Mol Immunol* 44 (9):2213-2220
181. Garcia-Laorden MI, Pena MJ, Caminero JA, Garcia-Saavedra A, Campos-Herrero MI, Caballero A, and Rodriguez-Gallego C (2006) Influence of mannose-binding lectin on HIV infection and tuberculosis in a Western-European population. *Mol Immunol* 43 (14):2143-2150
182. El Sahly HM, Reich RA, Dou SJ, Musser JM, and Graviss EA (2004) The effect of mannose binding lectin gene polymorphisms on susceptibility to tuberculosis in different ethnic groups. *Scand J Infect Dis* 36 (2):106-108
183. Bellamy R, Ruwende C, McAdam KP, Thursz M, Sumiya M, Summerfield J, Gilbert SC, Corrah T, Kwiatkowski D, Whittle HC, and Hill AV (1998) Mannose binding protein

- deficiency is not associated with malaria, hepatitis B carriage nor tuberculosis in Africans. *QJM* 91 (1):13-18
184. Druszczynska M, Strapagiel D, Kwiatkowska S, Kowalewicz-Kulbat M, Rozalska B, Chmiela M, and Rudnicka W (2006) Tuberculosis bacilli still posing a threat. Polymorphism of genes regulating anti-mycobacterial properties of macrophages. *Pol J Microbiol* 55 (1):7-12
 185. Capparelli R, Iannaccone M, Palumbo D, Medaglia C, Moscariello E, Russo A, and Iannelli D (2009) Role played by human mannose-binding lectin polymorphisms in pulmonary tuberculosis. *The Journal of infectious diseases* 199 (5):666-672
 186. Selvaraj P, Narayanan PR, and Reetha AM (1999) Association of functional mutant homozygotes of the mannose binding protein gene with susceptibility to pulmonary tuberculosis in India. *Tuber Lung Dis* 79 (4):221-227
 187. Liu W, Zhang F, Xin ZT, Zhao QM, Wu XM, Zhang PH, de VS, Richardus JH, Habbema JD, Yang H, and Cao WC (2006) Sequence variations in the MBL gene and their relationship to pulmonary tuberculosis in the Chinese Han population. *Int J Tuberc Lung Dis* 10 (10):1098-1103
 188. Gomez LM, Anaya JM, Vilchez JR, Cadena J, Hinojosa R, Velez L, Lopez-Nevot MA, and Martin J (2007) A polymorphism in the inducible nitric oxide synthase gene is associated with tuberculosis. *Tuberculosis (Edinb)* 87 (4):288-294
 189. Velez DR, Hulme WF, Myers JL, Weinberg JB, Levesque MC, Stryjewski ME, Abbate E, Estevan R, Patillo SG, Gilbert JR, Hamilton CD, and Scott WK (2009) NOS2A, TLR4, and IFNGR1 interactions influence pulmonary tuberculosis susceptibility in African-Americans. *Hum Genet* 126 (5):643-653
 190. Xiao J, Sun L, Jiao W, Li Z, Zhao S, Li H, Jin J, Jiao A, Guo Y, Jiang Z, Mokrousov I, and Shen A (2009) Lack of association between polymorphisms in the P2X7 gene and tuberculosis in a Chinese Han population. *FEMS Immunol Med Microbiol* 55 (1):107-111
 191. Li CM, Campbell SJ, Kumararatne DS, Bellamy R, Ruwende C, McAdam KP, Hill AV, and Lammas DA (2002) Association of a polymorphism in the P2X7 gene with tuberculosis in a Gambian population. *J Infect Dis* 186 (10):1458-1462
 192. Fernando SL, Saunders BM, Sluyter R, Skarratt KK, Goldberg H, Marks GB, Wiley JS, and Britton WJ (2007) A polymorphism in the P2X7 gene increases susceptibility to extrapulmonary tuberculosis. *Am J Respir Crit Care Med* 175 (4):360-366
 193. Nino-Moreno P, Portales-Perez D, Hernandez-Castro B, Portales-Cervantes L, Flores-Meraz V, Baranda L, Gomez-Gomez A, cuna-Alonzo V, Granados J, and Gonzalez-Amaro R (2007) P2X7 and NRAM1/SLC11 A1 gene polymorphisms in Mexican mestizo patients with pulmonary tuberculosis. *Clin Exp Immunol* 148 (3):469-477
 194. Mokrousov I, Sapozhnikova N, and Narvskaya O (2008) Mycobacterium tuberculosis co-existence with humans: making an imprint on the macrophage P2X7 receptor gene? *Journal of Medical Microbiology* 57:581-584
 195. Lamsyah H, Rueda B, Baassi L, Elaouad R, Bottini N, Sadki K, and Martin J (2009) Association of PTPN22 gene functional variants with development of pulmonary tuberculosis in Moroccan population. *Tissue Antigens* 74 (3):228-232

196. Malik S, Abel L, Tooker H, Poon A, Simkin L, Girard M, Adams GJ, Starke JR, Smith KC, Graviss EA, Musser JM, and Schurr E (2005) Alleles of the NRAMP1 gene are risk factors for pediatric tuberculosis disease. *Proc Natl Acad Sci U S A* 102 (34):12183-12188
197. Vejbaesya S, Chierakul N, Luangtrakool P, and Sermduangprateep C (2007) NRAMP1 and TNF-alpha polymorphisms and susceptibility to tuberculosis in Thais. *Respirology* 12 (2):202-206
198. El Baghdadi J, Remus N, Benslimane A, El Annaz H, Chentoufi M, Abel L, and Schurr E (2003) Variants of the human NRAMP1 gene and susceptibility to tuberculosis in Morocco. *Int J Tuberc Lung Dis* 7 (6):599-602
199. Leung KH, Yip SP, Wong WS, Yiu LS, Chan KK, Lai WM, Chow EY, Lin CK, Yam WC, and Chan KS (2007) Sex- and age-dependent association of SLC11A1 polymorphisms with tuberculosis in Chinese: a case control study. *BMC Infect Dis* 7:19
200. Ryu S, Park YK, Bai GH, Kim SJ, Park SN, and Kang S (2000) 3'UTR polymorphisms in the NRAMP1 gene are associated with susceptibility to tuberculosis in Koreans. *Int J Tuberc Lung Dis* 4 (6):577-580
201. Jin J, Sun L, Jiao W, Zhao S, Li H, Guan X, Jiao A, Jiang Z, and Shen A (2009) SLC11A1 (Formerly NRAMP1) gene polymorphisms associated with pediatric tuberculosis in China. *Clin Infect Dis* 48 (6):733-738
202. Awomoyi A, Sirugo G, Newport MJ, and Tishkoff S (2006) Global distribution of a novel trinucleotide microsatellite polymorphism (ATA)_n in intron 8 of the SLC11A1 gene and susceptibility to pulmonary tuberculosis. *Int J Immunogenet* 33 (1):11-15
203. Qu YB, Tang YX, Zhang ZB, Zhu R, Liu J, Gu SY, Lu GL, and Xia ZL (2006) [Relationship between single nucleotide polymorphisms of NRAMP1 gene and susceptibility to pulmonary tuberculosis in workers exposed to silica dusts]. *Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi* 24 (9):531-533
204. Delgado JC, Baena A, Thim S, and Goldfeld AE (2002) Ethnic-specific genetic associations with pulmonary tuberculosis. *J Infect Dis* 186 (10):1463-1468
205. Liu W, Zhang CY, Tian L, Li CZ, Wu XM, Zhao QM, Zhang PH, Yang SM, Yang H, and Cao WC (2003) [A case-control study on natural-resistance-associated macrophage protein 1 gene polymorphisms and susceptibility to pulmonary tuberculosis]. *Zhonghua Yu Fang Yi Xue Za Zhi* 37 (6):408-411
206. Zhang W, Shao L, Weng X, Hu Z, Jin A, Chen S, Pang M, and Chen ZW (2005) Variants of the natural resistance-associated macrophage protein 1 gene (NRAMP1) are associated with severe forms of pulmonary tuberculosis. *Clin Infect Dis* 40 (9):1232-1236
207. Liu W, Cao WC, Zhang CY, Tian L, Wu XM, Habbema JD, Zhao QM, Zhang PH, Xin ZT, Li CZ, and Yang H (2004) VDR and NRAMP1 gene polymorphisms in susceptibility to pulmonary tuberculosis among the Chinese Han population: a case-control study. *Int J Tuberc Lung Dis* 8 (4):428-434
208. Abe T, Iinuma Y, Ando M, Yokoyama T, Yamamoto T, Nakashima K, Takagi N, Baba H, Hasegawa Y, and Shimokata K (2003) NRAMP1 polymorphisms, susceptibility and clinical features of tuberculosis. *J Infect* 46 (4):215-220

209. Taype CA, Castro JC, Accinelli RA, Herrera-Velit P, Shaw MA, and Espinoza JR (2006) Association between SLC11A1 polymorphisms and susceptibility to different clinical forms of tuberculosis in the Peruvian population. *Infect Genet Evol* 6 (5):361-367
210. Awomoyi AA, Marchant A, Howson JM, McAdam KP, Blackwell JM, and Newport MJ (2002) Interleukin-10, polymorphism in SLC11A1 (formerly NRAMP1), and susceptibility to tuberculosis. *J Infect Dis* 186 (12):1808-1814
211. Ma X, Dou S, Wright JA, Reich RA, Teeter LD, El Sahly HM, Awe RJ, Musser JM, and Graviss EA (2002) 5' dinucleotide repeat polymorphism of NRAMP1 and susceptibility to tuberculosis among Caucasian patients in Houston, Texas. *Int J Tuberc Lung Dis* 6 (9):818-823
212. Gao PS, Fujishima S, Mao XQ, Remus N, Kanda M, Enomoto T, Dake Y, Bottini N, Tabuchi M, Hasegawa N, Yamaguchi K, Tiemessen C, Hopkin JM, Shirakawa T, and Kishi F (2000) Genetic variants of NRAMP1 and active tuberculosis in Japanese populations. International Tuberculosis Genetics Team. *Clin Genet* 58 (1):74-76
213. Puzyrev VP, Freidin MB, Rudko AA, Strelis AK, and Kolokolova OV (2002) [Polymorphisms of the candidate genes for genetic susceptibility to tuberculosis in the Slavic population of Siberia: a pilot study]. *Mol Biol (Mosk)* 36 (5):788-791
214. Liaw YS, Tsai-Wu JJ, Wu CH, Hung CC, Lee CN, Yang PC, Luh KT, and Kuo SH (2002) Variations in the NRAMP1 gene and susceptibility of tuberculosis in Taiwanese. *Int J Tuberc Lung Dis* 6 (5):454-460
215. Cervino AC, Lakiss S, Sow O, and Hill AV (2000) Allelic association between the NRAMP1 gene and susceptibility to tuberculosis in Guinea-Conakry. *Ann Hum Genet* 64 (Pt 6):507-512
216. Duan HF, Zhou XH, Ma Y, Li CY, Chen XY, Gao WW, and Zheng SH (2003) [A study on the association of 3'UTR polymorphisms of NRAMP1 gene with susceptibility to tuberculosis in Hans]. *Zhonghua Jie He He Hu Xi Za Zhi* 26 (5):286-289
217. Tosh K, Campbell SJ, Fielding K, Sillah J, Bah B, Gustafson P, Manneh K, Lisse I, Sirugo G, Bennett S, Aaby P, McAdam KP, Bah-Sow O, Lienhardt C, Kramnik I, and Hill AV (2006) Variants in the SP110 gene are associated with genetic susceptibility to tuberculosis in West Africa. *Proc Natl Acad Sci U S A* 103 (27):10364-10368
218. Babb C, Keet EH, van Helden PD, and Hoal EG (2007) SP110 polymorphisms are not associated with pulmonary tuberculosis in a South African population. *Hum Genet* 121 (3-4):521-522
219. Szeszko JS, Healy B, Stevens H, Balabanova Y, Drobniowski F, Todd JA, and Nejentsev S (2006) Resequencing and association analysis of the SP110 gene in adult pulmonary tuberculosis. *Hum Genet* 121 (2):155-160
220. Thye T, Browne EN, Chinbuah MA, Gyapong J, Osei I, Owusu-Dabo E, Niemann S, Rusch-Gerdes S, Horstmann RD, and Meyer CG (2006) No associations of human pulmonary tuberculosis with Sp110 variants. *J Med Genet* 43 (7):e32

221. Khor CC, Chapman SJ, Vannberg FO, Dunne A, Murphy C, Ling EY, Frodsham AJ et al (2007) A Mal functional variant is associated with protection against invasive pneumococcal disease, bacteremia, malaria and tuberculosis. *Nat Genet* 39 (4):523-528
222. Castiblanco J, Varela DC, Castano-Rodriguez N, Rojas-Villarraga A, Hincapie ME, and Anaya JM (2008) TIRAP (MAL) S180L polymorphism is a common protective factor against developing tuberculosis and systemic lupus erythematosus. *Infect Genet Evol* 8 (5):541-544
223. Nejentsev S, Thye T, Szeszko JS, Stevens H, Balabanova Y, Chinbuah AM, Hibberd M, Van D, V, Alisjahbana B, van CR, Ottenhoff TH, Png E, Drobniewski F, Todd JA, Seielstad M, and Horstmann RD (2008) Analysis of association of the TIRAP (MAL) S180L variant and tuberculosis in three populations. *Nat Genet* 40 (3):261-262
224. Hawn TR, Dunstan SJ, Thwaites GE, Simmons CP, Thuong NT, Lan NT, Quy HT, Chau TT, Hieu NT, Rodrigues S, Janer M, Zhao LP, Hien TT, Farrar JJ, and Aderem A (2006) A polymorphism in Toll-interleukin 1 receptor domain containing adaptor protein is associated with susceptibility to meningeal tuberculosis. *J Infect Dis* 194 (8):1127-1134
225. Ma X, Liu Y, Gowen BB, Graviss EA, Clark AG, and Musser JM (2007) Full-exon resequencing reveals toll-like receptor variants contribute to human susceptibility to tuberculosis disease. *PLoS ONE* 2 (12):e1318
226. Yim JJ, Lee HW, Lee HS, Kim YW, Han SK, Shim YS, and Holland SM (2006) The association between microsatellite polymorphisms in intron II of the human Toll-like receptor 2 gene and tuberculosis among Koreans. *Genes Immun* 7 (2):150-155
227. Ogus AC, Yoldas B, Ozdemir T, Uguz A, Olcen S, Keser I, Coskun M, Cilli A, and Yegin O (2004) The Arg753Gln polymorphism of the human toll-like receptor 2 gene in tuberculosis disease. *Eur Respir J* 23 (2):219-223
228. Velez DR, Wejse C, Stryjewski ME, Abbate E, Hulme WF, Myers JL, Estevan R, Patillo SG, Olesen R, Tacconelli A, Sirugo G, Gilbert JR, Hamilton CD, and Scott WK (2009) Variants in toll-like receptors 2 and 9 influence susceptibility to pulmonary tuberculosis in Caucasians, African-Americans, and West Africans. *Hum Genet* DOI: 10.1007/s00439-009-0741-7
229. Newport MJ, Allen A, Awomoyi AA, Dunstan SJ, McKinney E, Marchant A, and Sirugo G (2004) The toll-like receptor 4 Asp299Gly variant: no influence on LPS responsiveness or susceptibility to pulmonary tuberculosis in The Gambia. *Tuberculosis (Edinb)* 84 (6):347-352
230. Möller M, Flachsbarth F, Till A, Thye T, Horstmann RD, Meyer CG, Osei I, van Helden PD, Hoal E, Schreiber S, Nebel A, and Franke A (2010) A functional haplotype in the 3'UTR of *TNFRSF1B* is associated with TB in two African populations. *Am J Respir Crit Care Med* 181:388-393
231. Selvaraj P, Kurian SM, Chandra G, Reetha AM, Charles N, and Narayanan PR (2004) Vitamin D receptor gene variants of BsmI, ApaI, TaqI, and FokI polymorphisms in spinal tuberculosis. *Clin Genet* 65 (1):73-76
232. Wilkinson RJ, Llewelyn M, Toossi Z, Patel P, Pasvol G, Lalvani A, Wright D, Latif M, and Davidson RN (2000) Influence of vitamin D deficiency and vitamin D receptor

- polymorphisms on tuberculosis among Gujarati Asians in west London: a case-control study. *Lancet* 355 (9204):618-621
233. Liu W, Zhang CY, Wu XM, Tian L, Li CZ, Zhao QM, Zhang PH, Yang SM, Yang H, Zhang XT, and Cao WC (2003) [A case-control study on the vitamin D receptor gene polymorphisms and susceptibility to pulmonary tuberculosis]. *Zhonghua Liu Xing Bing Xue Za Zhi* 24 (5):389-392
 234. Roth DE, Soto G, Arenas F, Bautista CT, Ortiz J, Rodriguez R, Cabrera L, and Gilman RH (2004) Association between Vitamin D Receptor Gene Polymorphisms and Response to Treatment of Pulmonary Tuberculosis. *J Infect Dis* 190 (5):920-927
 235. Bornman L, Campbell SJ, Fielding K, Bah B, Sillah J, Gustafson P, Manneh K, Lisse I, Allen A, Sirugo G, Sylla A, Aaby P, McAdam KP, Bah-Sow O, Bennett S, Lienhardt C, and Hill AV (2004) Vitamin D receptor polymorphisms and susceptibility to tuberculosis in West Africa: a case-control and family study. *J Infect Dis* 190 (9):1631-1641
 236. Bellamy R, Ruwende C, Corrah T, McAdam KP, Thursz M, Whittle HC, and Hill AV (1999) Tuberculosis and chronic hepatitis B virus infection in Africans and variation in the vitamin D receptor gene. *J Infect Dis* 179 (3):721-724
 237. de WE, Delpont W, Rugamika CE, Meintjes A, Moller M, van Helden PD, Seoighe C, and Hoal EG (2010) Genome-wide analysis of the structure of the South African Coloured Population in the Western Cape. *Hum Genet* 128 (2):145-153
 238. Healy DG (2006) Case-control studies in the genomic era: a clinician's guide. *Lancet Neurol* 5 (8):701-707
 239. Barreiro LB, Neyrolles O, Babb CL, van Helden PD, Gicquel B, Hoal EG, and Quintana-Murci L (2007) Length Variation of DC-SIGN and L-SIGN Neck-Region has no Impact on Tuberculosis Susceptibility. *Human Immunology* 68 (2):106-112
 240. Western Cape Department of Health. TB Statistics: Cape Town. 2006.
 241. Kritzinger FE, den BS, Verver S, Enarson DA, Lombard CJ, Borgdorff MW, Gie RP, and Beyers N (2009) No decrease in annual risk of tuberculosis infection in endemic area in Cape Town, South Africa. *Trop Med Int Health* 14 (2):136-142
 242. Munch Z, Van Lill SWP, Booysen CN, Zietsman HL, Enarson DA, and Beyers N (2003) Tuberculosis transmission patterns in a high-incidence area: a spatial analysis. *International Journal of Tuberculosis and Lung Disease* 7 (3):271-277
 243. Holland PM, Abramson RD, Watson R, and Gelfand DH (1991) Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proc Natl Acad Sci U S A* 88 (16):7276-7280
 244. De la Vega FM, Lazaruk KD, Rhodes MD, and Wenz MH (2005) Assessment of two flexible and compatible SNP genotyping platforms: TaqMan SNP Genotyping Assays and the SNPlex Genotyping System. *Mutat Res* 573 (1-2):111-35
 245. Livak KJ (1999) Allelic discrimination using fluorogenic probes and the 5' nuclease assay. *Genet Anal* 14 (5-6):143-149

246. www.servicexs.com/plaatjes/TaqMan_AD_SNP_assay.jpg. TaqMan SNP Assay. 2010.
247. www.kapabiosystems.com/public/files/images/products/next-generation/kapa-probe/image1-modal.png. TaqMan Cluster Plot. 2010.
248. Hardy GH (2003) Mendelian proportions in a mixed population. 1908. *Yale J Biol Med* 76 (2):79-80
249. Weinberg W (1908) On the demonstration of heredity in man. *Naturkunde in Wurttemberg, Stuttgart* 64:368-382
250. Raven, Johnson, Losos, and Singer (2005) *Genes Within Populations*. Biology. Mc Graw-Hill Higher Education,
251. Xu J, Turner A, Little J, Bleecker ER, and Meyers DA (2002) Positive results in association studies are associated with departure from Hardy-Weinberg equilibrium: hint for genotyping error? *Hum Genet* 111 (6):573-4
252. Wigginton JE, Cutler DJ, and Abecasis GR (2005) A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet* 76 (5):887-93
253. Fisher RA (1922) On the interpretation of X^2 from contingency tables, and the calculation of P. *Royal Statistical Society* 85 (1):87-94
254. Dudbridge F (2003) Pedigree disequilibrium tests for multilocus haplotypes. *Genet Epidemiol* 25 (2):115-121
255. Barrett JC, Fry B, Maller J, and Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics (Oxford, England)* 21 (2):263-265
256. Cuervo AM, Gomes AV, Barnes JA, and Dice JF (2000) Selective Degredation of Annexins by Chaperone-mediated Autophagy. *The Journal of Biological Chemistry* 275 (43):33329-33335
257. Duncan R, Carpenter B, Main LC, Telfer C, and Murray GI (2008) Characterisation of protein expression profiling of annexins in colorectal cancer. *British Journal of Cancer* 98:426-433
258. Moss SE and Morgan RO (2004) The annexins. *Genome Biol* 5 (4):219
259. Futter CE and White IJ (2007) Annexins and endocytosis. *Traffic* 8 (8):951-958
260. Gruenberg J and Emans N (1993) Annexins in membrane traffic. *Trends Cell Biol* 3 (7):224-227
261. Gerke V, Creutz CE, and Moss SE (2005) Annexins: Linking Ca^{2+} Signalling to Membrane Dynamics. *Nature Reviews Molecular Cell Biology* 6:449-461
262. Rescher U and Gerke V (2004) Annexins - unique membrane binding proteins with diverse functions. *Journal of Cell Science* 117:2631-2639
263. Gerke V and Moss SE (2002) Annexins: from structure to function. *Physiol Rev* 82 (2):331-371

264. Hannon R, Croxtall JD, Getting SJ, Roviezzo F, Yona S, Paul-Clark MJ, Gavins FN, Perretti M, Morris JF, Buckingham JC, and Flower RJ (2003) Aberrant inflammation and resistance to glucocorticoids in annexin 1-/- mouse. *Federation of American Societies for Experimental Biology Journal* 17 (2):253-255
265. Srivastava M, Atwater I, Glasman M, Leighton X, Goping G, Caohuy H, Miller G, Pichel J, Westphal H, Mears D, Rojas E, and Pollard HB (1999) Defects in inositol 1,4,5-trisphosphate receptor expression, Ca(2+) signaling, and insulin secretion in the anx7(+/-) knockout mouse. *PNAS* 96 (24):13783-13788
266. Herr C, Smyth N, Ullrich S, Yun F, Sasse P, Hescheler J, Fleischmann B, Lasek K, Brixius K, Fässler R, Schröder R, and Noegel AA (2001) Loss of annexin A7 leads to alterations in frequency-induced shortening of isolated murine cardiomyocytes. *Molecular and Cellular Biology* 21 (13):4119-4128
267. Tomas A, Futter C, and Moss SE (2004) Annexin 11 is required for midbody formation and completion of the terminal phase of cytokinesis. *J Cell Biol* 165 (6):813-822
268. Lecona E, Turnay J, Olmo N, Guzman-Aranguez A, Morgan RO, Fernandez M-P, and Lizarbe MA (2003) Structural and functional characterization of recombinant mouse annexin A11: influence of calcium binding. *Biochem J* 373:437-449
269. Song J, Shih I, Chan DW, and Zhang Z (2009) Suppression of annexin A11 in ovarian cancer: implications in chemoresistance. *Neoplasia* 11 (6):605-14, 1
270. Bances P, Fernandez M-R, Rodriguez-Garcia M-I, Morgan RO, and Fernandez M-P (2000) Annexin A11 (ANXA11) Gene Structure as the Progenitor of Paralogous Annexins and Source of Orthologous cDNA Isoforms. *Genomics* 69:95-103
271. Tomas A and Moss SE (2003) Calcium- and cell cycle-dependent association of annexin 11 with the nuclear envelope. *J Biol Chem* 278:20210-20216
272. Paweletz CP, Ornstein DK, Roth MJ, Bichsel VE, Gillespie JW, Calvert VS, Vocke CD, Hewitt SM, Duray PH, Herring J, Wang QH, Hu N, Linehan WM, Taylor PR, Liotta LA, Emmert-Buck MR, and Petricoin EF3 (2000) Loss of annexin 1 correlates with early onset of tumorigenesis in esophageal and prostate carcinoma. *Cancer Research* 60 (60):6293-6297
273. Sharma MR, Koltowski L, Ownbey RT, Tuszyński GP, and Sharma MC (2006) Angiogenesis-associated protein annexin II in breast cancer: selective expression in invasive breast cancer and contribution to tumor invasion and progression. *Experimental and Molecular Pathology* 81 (2):146-156
274. Srivastava M, Bubendorf L, Srikantan V, Fossom L, Nolan L, Glasman M, Leighton X, Fehrl W, Pittaluga S, Raffeld M, Koivisto P, Willi N, Gasser TC, Kononen J, Sauter G, Kallioniemi OP, Srivastava S, and Pollard HB (2001) ANX7, a candidate tumor suppressor gene for prostate cancer. *PNAS* 98 (8):4575-4580
275. Xin W, Rhodes DR, Ingold C, Chinnaiyan AM, and Rubin MA (2003) Dysregulation of the annexin family protein family is associated with prostate cancer progression. *The American Journal of Pathology* 162 (1):255-261

276. Song J, Shih I, Salani R, Chan DW, and Zhang Z (2007) Annexin XI is associated with cisplatin resistance and related to tumor recurrence in ovarian cancer patients. *Clin Cancer Res* 13 (22 Pt 1):6842-6849
277. Hofmann S, Franke A, Fischer A, Jacobs G, Nothnagel M, Gaede KI, Schurmann M, Muller-Quernheim J, Krawczak M, Rosenstiel P, and Schreiber S (2008) Genome-wide association study identifies ANXA11 as a new susceptibility locus for sarcoidosis. *Nat Genet* 40 (9):1103-1106
278. Fabrellas EF (2007) Epidemiology of Sarcoidosis. *Arch Bronconeumol* 43 (2):92-100
279. Newman LS, Rose CS, and Maier LA (1997) Sarcoidosis. *N Engl J Med* 336 (17):1224-1234
280. American Thoracic Society (1999) Statement on sarcoidosis. Joint Statement of the American Thoracic Society (ATS), the European Respiratory Society (ERS) and the World Association of Sarcoidosis and Other Granulomatous Disorders (WASOG) adopted by the ATS Board of Directors and by the ERS Executive Committee, February 1999. *Am J Respir Crit Care Med* 160 (2):736-755
281. Iannuzzi MC (2004) Clara cell protein in sarcoidosis: another job for the respiratory tract protector? *Am J Respir Crit Care Med* 169 (2):143-144
282. McGrath DS, Goh N, Foley PJ, and du Bois RM (2001) Sarcoidosis: genes and microbes--soil or seed? *Sarcoidosis Vasc Diffuse Lung Dis* 18 (2):149-164
283. Noppen M (1999) Tuberculosis and sarcoidosis. *Chest* 115 (5):1480
284. Litinsky I, Elkayam O, Flusser G, Segal R, Yaron M, and Caspi D (2002) Sarcoidosis: TB or not TB? *Ann Rheum Dis* 61 (5):385-6
285. Nicodemus KK, Liu W, Chase GA, Tsai YY, and Fallin MD (2005) Comparison of type I error for multiple test corrections in large single-nucleotide polymorphism studies using principal components versus haplotype blocking algorithms. *BMC Genet* 6 Suppl 1:S78
286. Smith MW, Patterson N, Lautenberger JA, Truelove AL, McDonald GJ, Waliszewska A, Kessing BD et al (2004) A high-density admixture map for disease gene discovery in African Americans. *Am J Hum Genet* 74 (5):1001-13
287. Tishkoff SA, Reed FA, Friedlaender FR, Ehret C, Ranciaro A, Froment A, Hirbo JB, Awomoyi AA, Bodo JM, Doumbo O, Ibrahim M, Juma AT, Kotze MJ, Lema G, Moore JH, Mortensen H, Nyambo TB, Omar SA, Powell K, Pretorius GS, Smith MW, Thera MA, Wambebe C, Weber JL, and Williams SM (2009) The genetic structure and history of Africans and African Americans. *Science* 324 (5930):1035-1044
288. Mattick, J. Video Q&A: Non-coding RNAs and eukaryotic evolution - a personal view. 8(67). 2010. *BMC Biology*.
289. Bird A (2007) Perceptions of epigenetics. *Nature* 447 (7143):396-398
290. Bird A (2002) DNA methylation patterns and epigenetic memory. *Genes & Development* 16 (1):6-21

291. Holliday R and Pugh JE (1975) DNA modification mechanisms and gene activity during development. *Science* 187 (4173):226-232
292. Esteller M (2008) Epigenetics in Cancer. *The New England Journal of Medicine* 358 (11):1148-1159
293. Kornfeld H, Mancino G, and Colizzi V (1999) The role of macrophage cell death in tuberculosis. *Cell Death and Differentiation* 6:71-78
294. Lee J, Hartman M, and Kornfeld H (2009) Macrophage apoptosis in tuberculosis. *Yonsei Med J* 50 (1):1-11
295. Zychlinsky A (1993) Programmed cell death in infectious diseases. *Trends Microbiol* 1:114-117
296. Keane J, Balcewicz-Sablinska K, Remold HG, Chupp GL, Meek BB, Fenton MJ, and Kornfeld H (1997) Infection by *Mycobacterium tuberculosis* promotes human alveolar macrophage apoptosis. *Infect Immun* 65:298-304
297. Balcewicz-Sablinska MK, Keane J, Kornfeld H, and Remold HG (1998) Pathogenic *Mycobacterium tuberculosis* evades apoptosis of host macrophages by release of TNF-R2, resulting in inactivation of TNF-alpha. *J Immunol* 161 (5):2636-2641
298. Sly LM, Hingley-Wilson SM, Reiner NE, and McMaster WR (2003) Survival of *Mycobacterium tuberculosis* in host macrophages involves resistance to apoptosis dependent upon induction of antiapoptotic Bcl-2 family member Mcl-1. *Journal of immunology (Baltimore, Md : 1950)* 170:430-437
299. Spira A, Carrol JD, Liu G, Aziz Z, Shah V, and Kornfeld H (2003) Apoptosis genes in human alveolar macrophages infected with virulent or attenuated *Mycobacterium tuberculosis*: A pivotal role for tumor necrosis factor. *Am J Respir Cell Mol Biol* 29:545-551
300. Molloy A, Laochumroonvorapong P, and Kaplan G (1994) Apoptosis, but not necrosis, of infected monocytes is coupled with killing of intracellular bacillus Calmette-Guerin. *J Exp Med* 180:1499-1509
301. Oddo M, Renno T, Attinger A, Baker T, MacDonald HR, and Meylan PR (1998) Fas ligand-induced apoptosis of infected human macrophages reduces the viability on intracellular *Mycobacterium tuberculosis*. *Journal of immunology (Baltimore, Md : 1950)* 160:5448-5454
302. Riendeau CJ and Kornfeld H (2003) THP-1 cell apoptosis in response to *Mycobacterial* infection. *Infect Immun* 71:254-259
303. Velmurugan K, Chen B, Miller JL, Azogue S, Gurses S, and Hsu T (2007) *Mycobacterium tuberculosis nuoG* is a virulence gene that inhibits apoptosis of infected host cells. *PLoS Pathog* 3:e110
304. Fratazzi C, Arbeit RD, Carini C, and Remold HG (1997) Programmed cell death of *Mycobacterium avium* serovar 4-infected human macrophages prevents mycobacteria from spreading and induces mycobacterial growth inhibition by freshly added, uninfected macrophages. *Journal of immunology (Baltimore, Md : 1950)* 158:4320-4327

305. Lee J, Remold HG, leong MH, and Kornfeld H (2006) Macrophage apoptosis in response to high intracellular burden of *Mycobacterium tuberculosis* is mediated by a novel caspase-independent pathway. *J Immunol* 176 (7):4267-4274
306. Piacentini M and Colizzi V. <<Tissue>> transglutaminase: apoptosis versus autoimmunity. *Immunol Today* (in press)
307. Briken V and Miller JL (2008) Living on the edge: inhibition of host cell apoptosis by *Mycobacterium tuberculosis*. *Future Microbiol* 3:415-422
308. Gordon PB, Hoyvik H, and Seglen PO (1992) Prelysosomal and lysosomal connections between autophagy and endocytosis. *Biochem J* 283 (Pt 2):361-369
309. Jing X and Tang XM (1999) The convergent point of the endocytic and autophagic pathways in leydig cells. *Cell Research* 9:243-253
310. Liou W, Geuze HJ, Geelen MJ, and Slot JW (1997) The autophagic and endocytic pathways converge at the nascent autophagic vacuoles. *J Cell Biol* 136 (1):61-70
311. Deretic V, Singh S, Master S, Harris J, Roberts E, Kyei G, Davis A, de Haro S, Naylor J, Lee H-H, and Vergne I (2006) *Mycobacterium tuberculosis* inhibition of phagolysosome biogenesis and autophagy as a host defence mechanism. *Cellular Microbiology* 8 (5):719-727
312. Levine B (2005) Eating oneself and uninvited guests: autophagy-related pathways in cellular defense. *Cell* 120 (2):159-162
313. Gutierrez MG, Master SS, Singh SB, Taylor GA, Colombo MI, and Deretic V (2004) Autophagy is a defense mechanism inhibiting BCG and *Mycobacterium tuberculosis* survival in infected macrophages. *Cell* 119 (6):753-766
314. Ogawa M, Yoshimori T, Suzuki T, Sagara H, Mizushima N, and Sasakawa C (2005) Escape of intracellular *Shigella* from autophagy. *Science* 307 (5710):727-731
315. Paludan C, Schmid D, Landthaler M, Vockerodt M, Kube D, Tuschl T, and Munz C (2005) Endogenous MHC class II processing of a viral nuclear antigen after autophagy. *Science* 307 (5709):593-596
316. Masuda M, Yageta M, Fukuhara H, Kuramochi M, Maruyama T, Nomoto A, and Murakami Y (2002) The tumor suppressor protein TSLC1 is involved in cell-cell adhesion. *J Biol Chem* 277 (34):31014-31019
317. Boles KS, Barchet W, Diacovo T, Cella M, and Colonna M (2005) The tumor suppressor TSLC1/NECL-2 triggers NK-cell and CD8+ T-cell responses through the cell-surface receptor CRTAM. *Blood* 106 (3):779-786
318. Kakunaga S, Ikeda W, Itoh S, guchi-Tawarada M, Ohtsuka T, Mizoguchi A, and Takai Y (2005) Nectin-like molecule-1/TSL1/SynCAM3: a neural tissue-specific immunoglobulin-like cell-cell adhesion molecule localizing at non-junctional contact sites of presynaptic nerve terminals, axons and glia cell processes. *J Cell Sci* 118 (Pt 6):1267-1277
319. Shingai T, Ikeda W, Kakunaga S, Morimoto K, Takekuni K, Itoh S, Satoh K, Takeuchi M, Imai T, Monden M, and Takai Y (2003) Implications of nectin-like molecule-

- 2/IGSF4/RA175/SgIGSF/TSLC1/SynCAM1 in cell-cell adhesion and transmembrane protein localization in epithelial cells. *J Biol Chem* 278 (37):35421-35427
320. Biederer T, Sara Y, Mozhayeva M, Atasoy D, Liu X, Kavalali ET, and Sudhof TC (2002) SynCAM, a synaptic adhesion molecule that drives synapse assembly. *Science* 297 (5586):1525-1531
 321. Fukami T, Satoh H, Williams YN, Masuda M, Fukuhara H, Maruyama T, Yageta M, Kuramochi M, Takamoto S, and Murakami Y (2003) Isolation of the mouse Tsl1 and Tsl2 genes, orthologues of the human TSLC1-like genes 1 and 2 (TSL1 and TSL2). *Gene* 323:11-18
 322. Fukuhara H, Kuramochi M, Nobukuni T, Fukami T, Saino M, Maruyama T, Nomura S, Sekiya T, and Murakami Y (2001) Isolation of the TSL1 and TSL2 genes, members of the tumor suppressor TSLC1 gene family encoding transmembrane proteins. *Oncogene* 20 (38):5401-5407
 323. Galibert L, Diemer GS, Liu Z, Johnson RS, Smith JL, Walzer T, Comeau MR, Rauch CT, Wolfson MF, Sorensen RA, Van der Vuurst de Vries AR, Branstetter DG, Koelling RM, Scholler J, Fanslow WC, Baum PR, Derry JM, and Yan W (2005) Nectin-like protein 2 defines a subset of T-cell zone dendritic cells and is a ligand for class-I-restricted T-cell-associated molecule. *J Biol Chem* 280 (23):21955-21964
 324. Kuramochi M, Fukuhara H, Nobukuni T, Kanbe T, Maruyama T, Ghosh HP, Pletcher M, Isomura M, Onizuka M, Kitamura T, Sekiya T, Reeves RH, and Murakami Y (2001) TSLC1 is a tumor-suppressor gene in human non-small-cell lung cancer. *Nat Genet* 27 (4):427-430
 325. Murakami Y, Nobukuni T, Tamura K, Maruyama T, Sekiya T, Arai Y, Gomyou H, Tanigami A, Ohki M, Cabin D, Frischmeyer P, Hunt P, and Reeves RH (1998) Localization of tumor suppressor activity important in nonsmall cell lung carcinoma on chromosome 11q. *Proc Natl Acad Sci U S A* 95 (14):8153-8158
 326. Allinen M, Peri L, Kujala S, Lahti-Domenici J, Outila K, Karppinen SM, Launonen V, and Winqvist R (2002) Analysis of 11q21-24 loss of heterozygosity candidate target genes in breast cancer: indications of TSLC1 promoter hypermethylation. *Genes Chromosomes Cancer* 34 (4):384-389
 327. Jansen M, Fukushima N, Rosty C, Walter K, Altink R, Heek TV, Hruban R, Offerhaus JG, and Goggins M (2002) Aberrant methylation of the 5' CpG island of TSLC1 is common in pancreatic ductal adenocarcinoma and is first manifest in high-grade PanINs. *Cancer Biol Ther* 1 (3):293-296
 328. Saukkonen JJ, Bazydlo B, Thomas M, Strieter RM, Keane J, and Kornfeld H (2002) Beta-chemokines are induced by Mycobacterium tuberculosis and inhibit its growth. *Infect Immun* 70 (4):1684-1693
 329. Whittington HA, Armstrong L, Uppington KM, and Millar AB (2004) Interleukin-22: a potential immunomodulatory molecule in the lung. *Am J Respir Cell Mol Biol* 31 (2):220-226
 330. Wolk K, Witte E, Wallace E, Docke WD, Kunz S, Asadullah K, Volk HD, Sterry W, and Sabat R (2006) IL-22 regulates the expression of genes responsible for antimicrobial defense,

- cellular differentiation, and mobility in keratinocytes: a potential role in psoriasis. *Eur J Immunol* 36 (5):1309-1323
331. Zheng Y, Danilenko DM, Valdez P, Kasman I, Eastham-Anderson J, Wu J, and Ouyang W (2007) Interleukin-22, a T(H)17 cytokine, mediates IL-23-induced dermal inflammation and acanthosis. *Nature* 445 (7128):648-651
 332. Asefa B, Klarmann KD, Copeland NG, Gilbert DJ, Jenkins NA, and Keller JR (2004) The interferon-inducible p200 family of proteins: a perspective on their roles in cell cycle regulation and differentiation. *Blood Cells Mol Dis* 32 (1):155-167
 333. Peiper SC, Wang ZX, Neote K, Martin AW, Showell HJ, Conklyn MJ, Ogborne K, Hadley TJ, Lu ZH, Hesselgesser J, and Horuk R (1995) The Duffy antigen/receptor for chemokines (DARC) is expressed in endothelial cells of Duffy negative individuals who lack the erythrocyte receptor. *J Exp Med* 181 (4):1311-1317
 334. He W, Neil S, Kulkarni H, Wright E, Agan BK, Marconi VC, Dolan MJ, Weiss RA, and Ahuja SK (2008) Duffy antigen receptor for chemokines mediates trans-infection of HIV-1 from red blood cells to target cells and affects HIV-AIDS susceptibility. *Cell Host Microbe* 4 (1):52-62
 335. Davila S and Hibberd ML (2009) Genome-wide association studies are coming for human infectious diseases. *Genome Med* 1 (2):19
 336. de Bakker PIW and Telenti A. Infectious diseases not immune to genome-wide association. *Nature Genetics* (in press)

Addenda

ADDENDUM 1: BUFFERS, SOLUTIONS AND GELS

0.5M EDTA

93.06 g EDTA

400 ml distilled water

Adjust the pH to 8.0 using approximately 10g NaOH pellets

Make up to 500 ml with distilled water

Autoclave and store at room temperature

1M Tris

121g Tris

800ml distilled water

Adjust the pH to 8.0 with concentrated HCl

Make up to 1 litre with distilled water

Store at room temperature

2X Reducing SDS sample buffer

3.4 ml Tris (1M, pH 6.8)

2 ml Glycerol

3 ml of SDS (20%)

500µl Bromophenol blue

200µl EDTA (0.5M)

1 ml β-Mercaptoethanol

Store at -20 degrees

10% Ammonium persulphate (APS)

1 g APS

Make up to 10 ml with distilled water

Store at 4 degrees

10% Sodium Dodecyl Sulfate (SDS)

1 g SDS

Make up to 10 ml with distilled water

Store at room temperature

20X Sodium boride (SB) buffer

19.1 g di-sodium tetraborate

Make up to 500 ml with distilled water

Store at room temperature

20% Sodium Dodecyl Sulfate (SDS)

2 g SDS

Make up to 10 ml with distilled water

Store at room temperature

Agarose gel (1.5%)

3 g agarose
200 ml SB buffer (1X)

Blocking buffer

200 ml TBS-T wash buffer
10% fat free milk powder
1% BSA

Luria broth (LB)

10 g tryptone
5 g NaCl
5 g yeast extract
Make up to 1 litre with distilled water
Autoclave and store at room temperature

Running Buffer

3 g Tris
14.4 g Glycine
10 ml SDS (10%)
Make up to 1 litre with distilled water
Store at room temperature or 4 degrees

Separating gel (12%)

2.5 ml Acrylamide/bisacrylamide (40%)
2.5 ml Tris-HCl (1.5M, pH 8.8)
0.1 ml SDS (10%)
4.85 ml distilled water
10 µl TEMED
50 µl APS (10%)

Stacking gel (3%)

0.4 ml Acrylamide/bisacrylamide (40%)
0.75 ml Tris-HCl (1M, pH 6.8)
50 µl SDS (10%)
4 ml distilled water
10 µl TEMED
25 µl APS (10%)

TBS-T Wash Buffer

10 ml Tris (1M pH 7.6)
8 g NaCl
1 ml Tween 20
Make up to 1 litre with distilled water
Store at room temperature or 4 degrees

TE

1.21 g Tris-HCl

0.372 g EDTA

Add 800 ml distilled water

Adjust the pH to 8.0 with concentrated HCl

Make up to 1 litre with distilled water

Autoclave and store at room temperature

Transfer buffer

3.03 g Tris

14.4 g Glycine

200 ml methanol

Make up to 1 litre with distilled water

Store at 4 degrees

ADDENDUM 2: REAGENTS

Reagent	Supplier	
10-250 kDa PageRuler Prestained protein ladder	Fermentas	Canada
10X PCR buffer with 1.5mM MgCl ₂	Southern Cross Biotechnologies	South Africa
100 bp DNA ladder	Whitehead Scientific	South Africa
Acetone	Sigma-Aldrich	South Africa
Acrylamide/bisacrylamide (40%)	Sigma-Aldrich	South Africa
Agarose	Whitehead Scientific	South Africa
Ammonium persulphate (APS)	Sigma-Aldrich	South Africa
Antibody_primary	GeneTex®, Inc.	USA
Antibody_secondary	Santa Cruz Biotechnology Inc.	USA
BCG	Division of Molecular Biology and Human Genetics, Stellenbosch University	South Africa
β-Mercaptoethanol	Sigma-Aldrich	South Africa
Bioline dNTPs (4x 25μmol solutions)	Celtic Diagnostics	South Africa
Bromophenol blue	Sigma-Aldrich	South Africa
BSA	Sigma-Aldrich	South Africa
Chloroform	Sigma-Aldrich	South Africa
DEPC water	Ambion	USA
D _c Protein Assay	Bio-Rad	South Africa
E.coli FastMedia™ LB Agar Amp IPTG/X-Gal	Fermentas	Canada
E.coli JM109	Promega	USA
ECl detection fluid	Amersham Biosciences	UK
EDTA	Sigma-Aldrich	South Africa
Ethanol	Merck	South Africa
Ethidium bromide (10 mg/ml)	Sigma-Aldrich	South Africa
ExoSAPit	USB Corporation	USA

Reagent

Fat free milk powder
GIBCO® GlutaMAX™ media
Glycerol
Glycine
Histopaque H8889
Human AB Serum (10%)
Imprint™ DNA Modification Kit
Isopropanol
LightCycler® FastStart DNA Master SYBR Green I Kit
Nucleon BACC3 Kit for blood and cell cultures
Orange loading dye solution
Methanol
Middlebrook oleic-albumin-dextrose-catalase (OADC)
Middlebrook 7H9 broth
Middlebrook 7H11 Agar
PCR primers
pGEM-T Easy vector
Phenol
Promega 10X ligase buffer
Promega PureYield Plasmid Midiprep System
Promega T4 DNA ligase
Promega Wizard Plus SV Miniprep Kit
Promega Wizard SV Gel and PCR Clean-up System
RPMI-1640
Sodium Dodecyl Sulphate (SDS)
Sodium Chloride (NaCl)
Shrimp alkaline phosphatase

Supplier

Pick 'n Pay
Invitrogen
Sigma-Aldrich
Sigma-Aldrich
Sigma-Aldrich
Sigma-Aldrich
Sigma-Aldrich
Sigma-Aldrich
Roche
Amersham Biosciences
Inqaba Biotechnology
Sigma-Aldrich
BD Scientific
BD Scientific
BD Scientific
UCT
Promega
Sigma-Aldrich
Whitehead Scientific
Whitehead Scientific
Whitehead Scientific
Whitehead Scientific
Whitehead Scientific
Sigma-Aldrich
Sigma-Aldrich
Sigma-Aldrich
Amersham Biosciences

South Africa
USA
South Africa
South Africa
South Africa
South Africa
South Africa
South Africa
South Africa
UK
South Africa
South Africa
USA
USA
USA
South Africa
USA
South Africa
South Africa
South Africa
South Africa
South Africa
South Africa
South Africa
South Africa
South Africa
Germany

Reagent

Super-therm GOLD Hotstart TAQ DNA polymerase
T4 DNA ligase
TaqMan SNP Genotyping Assay Mix
TaqMan SNP Genotyping assays
TEMED
Transcriptor First Strand cDNA Synthesis Kit
Tris
TRIzol® reagent
Tween 20
Wizard Plus SV Miniprep Kit
Wizard SV Gel and PCR Clean-Up System

Supplier

Southern Cross Biotechnology
Promega
Applied Biosystems
Applied Biosystems
Sigma-Aldrich
Roche
Sigma-Aldrich
Invitrogen
Sigma-Aldrich
Promega
Promega

South Africa
USA
South Africa
South Africa
South Africa
Germany
South Africa
USA
South Africa
USA
USA

ADDENDUM 3: EQUIPMENT

Equipment	Supplier	
1ml/cc insulin syringe	Supra Latex	South Africa
6 well tissue culture dishes	Nunc	Denmark
96 deepwell storage plates	Eppendorf	Germany
ABI Prism 9700 HT Sequence Detection System	Applied Biosystems	Germany
Autoflow CO2 water jacketed incubator	Thermofisher Scientific	USA
BD Falcon tubes (15ml and 50ml)	Scientific Group	South Africa
Cell scraper	Corning Incorporated	USA
Consort electrophoresis power supply	Sigma-Aldrich	South Africa
Disposable serological pipettes (10ml)	Corning Incorporated	USA
Easypet	Eppendorf	Germany
EpMotion 5070	Eppendorf	Germany
EpTips Motion Filter 1-50ul	Merck	South Africa
Eppendorf 0.5, 1.5 and 2 ml tubes	Merck	South Africa
Eppendorf Centrifuge 5810R	Merck	South Africa
Eppendorf Mastercycler	Merck	South Africa
G-box	Syngene	UK
GeneAmp PCR system 9700	Applied Biosystems	Germany
Hybond-P PVDF	GE Healthcare	USA
Hypercasette™	Amersham Life Sciences	England
Hyper Processor	Amersham Pharmacia Biotech	USA
LightCycler 1.5 system	Roche	Germany
LightCycler® Capillary	Roche	Germany
Magnetic stirrer	Velp Scientifica	Italy
Medical X-ray film	Agfa	South Africa
MicroAmp 384-well reaction plates	Applied Biosystems	South Africa

Equipment**Supplier**

MicroAmp™ Optical Adhesive Covers milli-PORE	Applied Biosystems milli-PORE	South Africa USA
Nanodrop ND-1000 Spectrophotometer	Inqaba Biotechnology	South Africa
Orbital Shaker Incubator LM-530	Yih Der	Japan
Petri dishes	Nunc	Denmark
Pipettes (2µl, 20µl, 200µl, 1000µl)	Eppendorf	Germany
Pipette tips	Whitehead Scientific	South Africa
Protein gel apparatus	Bio-Rad	South Africa
qPCR capillaries	Roche	South Africa
Single and multi-channel pipettes	Eppendorf	Germany
Ultrospec 4051	LKB Biochrom	UK
Vacurette Heparin tubes	Greiner Bio-One	Austria
Vortex mixer VM-1000	Digisystem lab. Instruments, inc.	Taiwan
Western blot apparatus	Bio-Rad	South Africa
Western blotting filter paper	Thermo Scientific	USA

ADDENDUM 4: SOFTWARE PACKAGES

Bioedit Sequence Alignment Editor v7.05	Ibis Therapeutics, USA
Cocaphase	MRC Human Genome Mapping Project, UK
Epi Info 2000	Centres for Disease Control and Prevention, USA
Haploview v4.1.	Daly lab, Broad Institute, USA
Nanodrop v3.0.1	Inqaba Biotechnology, South Africa
Prism 5.02	GraphPad Software, Inc., California, USA
Sequencher v4.2	Gene Codes Corporation, USA
SDS 2.0	Applied Biosystems, Germany

ADDENDUM 5: ONLINE RESOURCES

BiQ Analyzer software package	biq-analyzer.bioinf.mpi-inf.mpg.de/
BLAST	www.ncbi.nlm.nih.gov/BLAST
CHIP Bioinformatics	snpper.chip.org/bio/snpper-enter
dbSNP	www.ncbi.nlm.nih.gov/SNP
HapMap	www.hapmap.org
Human BLAT search	genome.ucsc.edu/cgi-bin/hgBlat
PubMed	www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed